Speech Science Research Centre

Non-normative preaspirated voiceless fricatives in Scottish English: Phonetic and phonological characteristics

Olga B Gordeeva & James M Scobbie

Working Paper WP-12

January 2007

Queen Margaret University

EDINBURGH

QMU Speech Science Research Centre Working Paper WP12 (2007)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

Update and Sourcing Information 2013

This updated version with new front-matter is available online in pdf format

- 2013 onwards at http://eresearch.qmu.ac.uk/154/
- 2013 onwards at http://www.qmu.ac.uk/casl

Author Contact details:

- 2013 onwards at ogordeeva⊗gmail.com (@ not ⊗)

Subsequent publication & presentation details:

- This is the draft of a paper submitted for peer-review (2007)
- This is an early version of the paper published (2010) as Gordeeva, Olga B. and Scobbie, James M (2010) Preaspiration as a correlate of word-final voice in Scottish English fricatives. In: Fuchs, S., Toda, M. and Żygis, M. (Eds) *Turbulent Sounds: An Interdisciplinary Guide*. Mouton de Gruyter, Berlin, pp. 167-207. ISBN 978-3-11-022657-7 with QM repository information (including abstract) at http://eresearch.qmu.ac.uk/1605/
  The authors request that you consult and cite the published, peer-reviewed and amended version, not this working version.

QMU Speech Science Research Centre Working Paper WP12 (2007)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

QMU Speech Science Research Centre Working Paper WP12 (2007)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

QMU Speech Science Research Centre Working Paper WP12 (2007)
Series Editors: James M Scobbie, Ineke Mennen, Jocelynne Watson

Non-normative preaspirated voiceless fricatives in Scottish English: Phonetic and phonological characteristics

Olga B. Gordeeva and James M. Scobbie

Abstract

Preaspiration is usually associated with stops rather than fricatives, both at phonological and phonetic levels of description. This study reports the occurrence of phonetic (non-normative) preaspiration of voiceless fricatives in Scottish Standard English (SSE) spoken in the Central Belt of Scotland. We classify it as non-normative because it is variably present in different speakers, but the distribution is nevertheless understandable on phonetic grounds. The paper analyses the phonetic distribution of preaspiration and its functions in SSE.

Preaspiration is shown to occur more frequently after open vowels and phrase-finally. Sociophonetic conditioning by speaker's sex is not found to be relevant. Functional analysis shows that preaspiration (reflected in the amount of noise in mid/high spectral frequencies) is a systematic correlate of phonological fricative /voice/ contrast phrase-finally. In this context, it appears to be even stronger predictor of /±voice/ than such traditionally-considered correlates as voicing offset and segmental duration. The results show that abstract non-neutralised /voice/ is phonetically multidimensional such that fricative preaspiration can maintain the contrast in the contexts where phonetic voicing is demoted. The extent and functioning of preaspiration in SSE suggests that it is a variety-specific optional characteristic resulting from a learned dissociation of lingual and laryngeal stricture gestures in voiceless fricatives.

1. Introduction

1.1 Preaspiration

Preaspiration has been described as a co-ordinatory relationship between a vowel and a following voiceless segment (Laver, 1994). This involves an early offset of modal voicing in the vowel in anticipation of the wide opening of the vocal folds required by the voiceless segment (Ladefoged & Maddieson, 1996). The glottal opening is accompanied by variable in strength supraglottal turbulence that is often interchangeably termed as 'breathiness', 'aspiration' or 'whisper' (Laver, 1994, p.189-190). We report here for the first time in print the existence of preaspirated fricatives in the Standard English spoken in the Lowland Central Belt of Scotland (SSE). Preaspiration in this paper is viewed as varying in the precise combination of supraglottal constriction and glottal opening necessary to maintain turbulent flow, and, being gradient, could fall under any of the above terms.

Across languages and varieties, preaspiration can be 'normative' (a term introduced by Helgason, 2002): i.e. a consistent characteristic, as in Icelandic (Thráinsson, 1979), where it seems to be a major correlate of phonological contrasts exhibited by pairs like [viht] ("wide") vs. [vitt] ("breadth"). Icelandic preaspiration is part of the phonetic/phonological sound system, diachronically replacing an earlier contrast of gemination, and has a near-segmental status in the language.

Preaspiration can also be 'non-normative': i.e. variably present/absent in different speakers of a variety. This is the case in various Swedish and Norwegian dialects (van Dommelen, 1998; Helgason, 2002; Schaeffler, 2005), where this phonetic characteristic often complements phonological contrasts between short/long vowels and consonants. Similarly, preaspiration is an important perceptual cue to phonological /±voice/ in a trochee-medial context in words like *lake* ("brine") and *lage* ("to make") in some Norwegian varieties (van Dommelen, 1998). An intermediate situation is when preaspiration is socially-structured, as it is, for example, in Tyneside English (Docherty & Foulkes, 1999) where some young working class females use preaspirated word-final stops while such variants are virtually absent from the 45-65 age group (Docherty et al., 1999). In these latter ('non-normative') situations, preaspiration is not seen as a replacement of an earlier aspect of the sound system in the same way as Icelandic, but part of a synchronic phonetic continuum.

Only some languages are uncontroversially labeled as having normative preaspiration. For example, according to UPSID it only occurs in two languages – Icelandic and Scottish Gaelic, prompting claims that it is rare (Bladon, 1986; Silverman, 2003). Although low intensity aspiration following full vowels is prone to masking effects from the human auditory system that negatively influence its perception (Bladon, 1986), such auditory limitations are not sufficient to exclude preaspiration from the world's phonological inventories. On the contrary, speakers of the languages with /VhC/ sequences, such as Turkish and Arabic (where /h/ is a phoneme rather than obligatory intersegmental transition) are better attuned to the presence of [h] in such sequences than speakers of the languages, such as English or French, lacking segmental /VhC/ phonotactics (Mielke, 2003). These considerations make close phonetic examination of comparable or more arguable cases of language varieties with unstable non-normative preaspiration a very worthwhile research strategy, especially if the phonetic underpinnings of this phenomenon and its functions are to be understood, as well as more general processes of co-articulation, synchronic variation and diachronic change.

Most accounts of 'non-normative' preaspiration have so far considered preaspirated stops (e.g. van Dommelen, 1998; Helgason, 2002; Schaeffler, 2005). While there are phonological contrasts with word-initial postaspirated voiceless fricatives reported in languages like Burmese (see discussion in Vaux, 1998), there have been no reports of lexically-contrastive preaspirated fricatives in which preaspiration is clearly normative. Although the possibility of preaspirated fricatives in *non-normative* forms has been acknowledged previously, albeit in a very limited number of studies (Helgason, 2002; Jones & Llamas, 2003), there are no explanations proposed for its possible functions in speech. In particular, it is possible that it could enhance phonological contrasts (e.g. consonant /voice/ or vowel length), or convey sociolinguistic meaning in a way similar to stops, rather than just being a result of automatic co-articulation resulting from aerodynamics of the vocal tract as proposed by Gobl & Ní Chasaide (1999). The broader aims of this study are to bridge the gap between possible functional and co-articulatory explanations, to provide a synchronic account of the acoustic characteristics of preaspirated fricatives in the Standard English spoken in Scotland, and to develop an analytic method able to analyse the whole range of relative voicing offsets in vowel-fricative transitions.

The specific aims are:

(1) to account for the phonetic conditioning of preaspirated fricatives with regard to the influence of vowel height, sentence prosody and speaker sex;

(2) to quantitatively determine which acoustic properties are most important in differentiating preaspirated and non-preaspirated CVC words ending with voiceless fricatives (CVf);

(3) to investigate whether preaspiration of voiceless fricatives might function in SSE to prevent neutralisation of the phonological /voice/ contrast in specific phrasal contexts.

1.2 Scottish Standard English and the areal distribution of preaspiration

SSE spoken in the Lowland Central Belt has different preaspirating languages and varieties as geographical neighbours. There is phonological preaspiration of stops in some varieties of Scottish Gaelic (Ladefoged, Ladefoged, Turk, Hind, & Skilton St.John, 1999), and socio-phonetically structured non-normative preaspiration of stops in Tyneside English (Docherty et al., 1999; Watt & Allen, 2003). As opposed to that, there have been no reports of wide and early glottal abduction before word-final stops in SSE that could result in preaspiration. On the contrary, SSE stops are often glottalised (Wells, 1982; Stuart-Smith, 1999; Chirrey, 1999) or even produced with complete glottal closure as strong ejectives (Gordeeva & Scobbie, 2006).

Despite the seeming lack of preaspiration before stops, SSE female speakers have been noted to often produce word-final fricatives with substantial preaspiration (Gordeeva & Scobbie, 2004). That study defined preaspiration as whispery or weakly glottal Vf transitions longer than a threshold of 30 ms. Although preaspiration was variable in frequency of occurrence, it was observed in all five female MC speakers, and in 41% of all tokens (out of a total *n*=300). One speaker used them almost exclusively. They were more frequent in "bus" (74%) with a more open vowel than in "fish" (29%) or "goose" (21%) tokens. In terms of duration, in the more open vowels, preaspirated transitions could be as long as the vowel itself. Phrase-final location of target words increased the frequency of occurrence and yielded longer duration of preaspiration.

Figure 1 shows an example of a preaspirated transition produced by a female middle class (MC) speaker from Edinburgh in phrase-final "grass". The whispery to weakly aspirated Vf transition is 118 ms long, while the preceding full vowel is 129 ms.

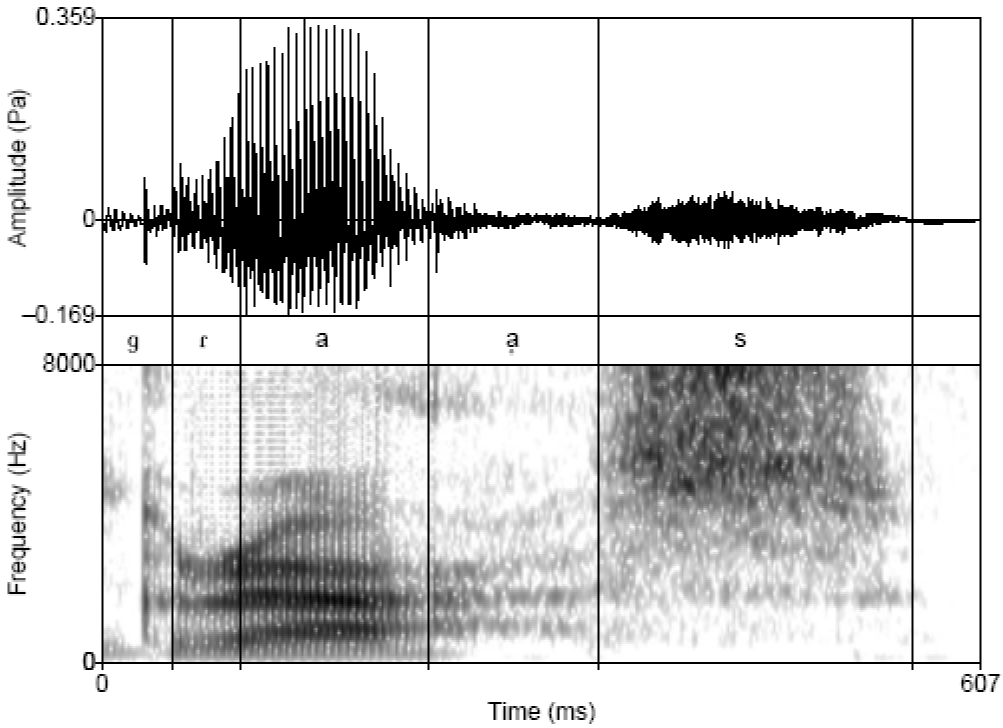


Figure 1. An example of whispery Vf transition in phrase-final "grass" token produced by a female MC speaker from Edinburgh.

Such extensive transitions after (mid-)open SSE vowels were often realised as devoiced vowels (see Figure 1) or as glottal [h]. For more narrowly constricted (mid-)close vowels, the airflow through a wide glottal opening and simultaneous supraglottal narrowing seemed to create additional turbulence noise at the palatal constriction resulting in palatal [ç]. This phenomenon is also known from the literature on (post-)aspiration (Kim, 1970; Stevens, 1998, p.445), as well as phonemic /h/ before /i/ and /j/ is often [ç] in various languages.

Although the extent of preaspiration shown in Figure 1 is clearly problematic for segmental annotation of vowel offsets and their acoustic analysis, previous phonetic studies that looked into vowel duration in SSE vowel-obstruent sequences (e.g. Agutter, 1988; McKenna, 1988) made no mention of extensively preaspirated fricatives. Indeed, McKenna's discussion of segmentation criteria noted the problems arising from the partial devoicing of /z/ and voicing of /s/, but does not mention preaspiration. The lack of other specifically Scottish English reports on preaspirated fricatives so far raises a question (further beyond the scope of this paper) whether we are observing an ongoing diachronical change in the gestural coordination of SSE vowel/voiceless fricative sequences. Recent large-scale socially-stratified phonetic studies of Glasgow English (summarised in Stuart-Smith, 2004) have focused on continuous speech and single-item wordlists, materials which may not favour the phenomenon, as we will see, so it is possible that it is a long-standing but previously unreported characteristic.

1.3 Voicing offset in vowel-fricative sequences

From a phonetic perspective, we could a priori expect to find preaspiration in fricatives. It is known that the laryngeal control mechanisms before voiceless fricatives seem to create better aerodynamic conditions for preaspiration to occur compared to the contexts before voiceless stops. The glottal abduction is initiated earlier in fricatives (than in stops) relative to the formation of the supraglottal constriction (Löfqvist & McGarr, 1987; Hoole, 1999); and it involves a substantially greater amplitude of vocal fold abduction (Löfqvist et al., 1987). Exactly these two conditions have also been shown to influence phonological /voice/ and postaspiration in word-initial stops (Lisker & Abramson, 1964; Kim, 1970).

It has been previously hypothesised that early glottal abduction (relative to supraglottal constriction) and a gradual breathy vowel offset might be an automatic and possibly universal feature of voiceless fricatives (Gobl & Ní Chasaide, 1999). The study found an earlier onset of glottal abduction in voiceless (compared to voiced) fricatives irrespective of the languages considered (Swedish, Italian and German). These languages showed striking similarity in the vowel source characteristics before voiceless fricatives: i.e. gradually falling excitation strength, gradually rising dynamic leakage and increasingly symmetrical shape of the glottal pulse: i.e. all signs of early glottal abduction and increasing breathiness.

Despite these arguments, cross-linguistically there are surprisingly few reports of aspiration in vowel/voiceless fricative transitions of the extent similar to that reported for preaspirated stops. So far, there have been notes about the possibility of preaspirated fricatives as a language-specific characteristic in the context of studies of languages/varieties with non-normative preaspiration of *stops*: i.e. it has been noted in Central Standard Swedish (Helgason, 2002) and reported in Middlesbrough variety of British English (Jones et al., 2003). Jones and Llamas's study on the characteristics of word-final fricated and preaspirated stops in Middlesbrough was based on the data of three speakers. The materials included the word "mat" compared to control fricatives in words like "mass" and "mash". The authors concluded that the duration and auditory quality of breathy offsets in vowel-voiceless fricative (Vf) transitions in the control items was so substantial that it could be labeled as 'preaspirated' in itself, and was comparable to preaspirated "mat" tokens. The mean preaspiration ratio in a word like "mass" was 0.54, meaning that preaspiration was about as long as the vowel itself.

Although it can be accepted as previously suggested (Gobl et al., 1999) that earlier glottal abduction before voiceless (compared to voiced) fricatives can be a cross-linguistic characteristic, the general lack of cross-linguistic phonetic reports of preaspiration of such large extent as in Middlesbrough and SSE casts serious doubt on the aerodynamic automaticity of extensive preaspiration resulting from this early abduction.

This reasoning implies that creating aspirated turbulence in Vf-transitions must be *learnt* either in terms of timing (Browman & Goldstein, 1992) and/or in terms of the abduction amplitude (Kim, 1970).

Uncovering variety-specific learnability and linguistic functions of large timing dissociations of laryngeal and oral gestures before voiceless fricatives has implications for theories of phonological representations viewed as gestures (Browman et al., 1992) and their acquisition. The variable dissociation could be learnable in a variety-specific way, with a possibility of ongoing sound changes (and possibly phonologisation) occurring in e.g. places where listener is prone to make mistakes in the interpretation of underlying forms (Ohala, 1993). In this case, preaspiration can be expected to assist (or even take over) functions in speech communication beyond mere free variation in linguistic, sociolinguistic or paralinguistic domains.

1.4 Functions of preaspiration in SSE

Assuming that the extent of preaspiration in SSE is variety-specific, what functions could it fulfill?

A sociophonetic function is one option, since our previous study (Gordeeva et al., 2004) found it in SSE female speakers, but had no male subjects. If it is females who predominately produce aspirated Vf-transitions, it might be a favored context for the sociolinguistic use of breathiness – a possibility strongly suggested by the data of Henton and Bladon (1985) for female speakers of several British English varieties other than SSE. (Other sociolinguistic functions related to class/age stratification are possible but will be outside the scope of this paper.)

The main function that we would like to address in this study is that of a simultaneous cue to the phonological fricative /voice/ contrast in particular prosodic contexts. There is, namely, an interesting parallel between the frequent lenition of voicing as a phonetic correlate of phonological /voice/ in British English in phrase-final positions (Haggard, 1978; Docherty, 1992) and the promotion of transitional aspiration in this context found in SSE (Gordeeva et al., 2004).

It is well known that depending on the phonetic context, phonological /voice/ of word-final obstruents in English can be controlled by a multitude of cues with voiceless fricatives showing much earlier cessation of voicing (Haggard, 1978; Docherty, 1992; Smith, 1997), longer consonantal and shorter vowel duration (Smith, 1997), higher voice source airflow during the consonant (Smith, 1997), and accordingly higher frication noise amplitude (Balise & Diehl, 1994). Most of these studies (apart from Smith, 1997, that also looked at vowel duration) analyzed the acoustic correlates of /voice/ within the fricative scope and did not include the preceding vowel. However, as we saw above, preaspiration in Vf sequences is an anticipatory event with the glottal abduction gesture potentially starting quite early in the vowel. Therefore, in order to find out whether it can be promoted as phonetic correlate of phonological /voice/ phrase-finally, we must increase the scope of preaspiration/voicing analysis to the *whole Vf-rhyme*.

Preaspiration of phrase-final voiceless fricatives may well be associated with phonetic devoicing of /+voice/ fricatives in this context. In SSE, unlike other Germanic languages, there is no phonological neutralisation. A priori this is because devoicing is incomplete, but perhaps the preaspiration is present as a compensatory measure, maintaining the contrast as in some Norwegian varieties (van Dommelen, 1998). If the "same" contrast between /+voice/ and /-voice/ fricatives is expressed across a range of different areas of phonetic space, then it should be evident in interspeaker variation. Such patterns were shown for Shetlandic English in the word-initial /voicing/ contrast in stops (Scobbie, 2006) for varying amounts of postaspiration and frequency of pre-voicing.

Since Gordeeva & Scobbie (2004) provided no statistical tests of the prosodic influences on the prevalence of preaspiration phrase-finally, this study will perform statistical tests of the frequency distributions of preaspirated and non-preaspirated fricatives in phrase-final and non-final contexts. This syntagmatic consideration of the appearance of preaspiration at prosodic edges is also relevant for the theoretical discussion about the mutual influences of segmental and prosodic levels of speech in areas such as domain-final lengthening and strengthening (Edwards, Beckman, & Fletcher, 1991; Fougeron & Keating, 1997; Cho, 2001) and ultimately of mental representation(s) of these levels (Keating & Shattuck-Hufnagel, 2002; Levelt, 1989).

1.5 Acoustic measures of (pre)aspiration and breathiness

With the conception of a /voice/ contrast (whether stable or not) being analysable in different regions of multidimensional phonetic space, we need a broader range of phonetic measures related to preaspiration in addition to previously used voicing offset, duration or frication amplitude as cues to /voice/.

One way to extend the range would be to use direct laryngeal techniques such as fiberoptic filming, transillumination or photoelectric glottography (e.g. Löfqvist & Yoshioka, 1980; Ní Chasaide, 1987) that have proven to be very useful in understanding the glottal abduction mechanisms behind aspiration. However, such experiments usually concern single case studies due to the procedural challenges imposed by the techniques.

On the other hand, there are acoustic methods for measuring glottal characteristics that do not require specialist equipment and allow processing bigger samples of subjects (Hillenbrand, Cleveland Ronald A, & Ericson Robert L, 1994; Holmberg, Hillman Robert E, Perkell Joseph S, Guiod Peter C, & Goldman Susan L, 1995; Hanson, 1997). Acoustic studies of breathiness suggest that its most robust predictors are the amount of noise present in mid- and high spectral frequencies (Klatt & Klatt, 1990; Hillenbrand et al., 1994) and the amplitude of the first (H1) relative to the second harmonic (H2) (Hillenbrand et al., 1994; Holmberg et al., 1995; Klatt et al., 1990).

Klatt and Klatt (1990) studied production and perception of long term vowel breathiness in 10 male and 10 female speakers of English. Although they found that breathy phonation was characterised by increased open quotient (high H1 relative to H2), the perception of breathiness was mainly correlated to increases of aspiration noise in the higher harmonics (around F3). In that study, the presence of higher-frequency aspiration was judged by manual examination of time-domain waveforms in syllables like [ha] and [ʔa]. Since unfiltered breathy signals often contained waving resulting form prominent low frequency component (H1) (see also Stevens, 1998, p. 425), the signal was band-pass filtered around F3 to allow the visual inspection. The presence of high frequency noise accounted for 60% of variance in the listeners' perception of long term breathiness in the syllables.

Hillenbrand at al (1994) study of production and perception of breathy vowel quality produced by male and female subjects confirmed the results of Klatt and Klatt

(1990). The high frequency noise accounted for about 80% of breathiness ratings in perception, while the amplitude of the first harmonic was the second best predictor.

To determine the amount of high(/mid) frequent noise in the vowels, Klatt and Klatt (1990) performed manual inspection of band-pass filtered waveforms. Hillenbrand et al. (1994) used automatic acoustic measures based on cepstral and pitch autocorrelation algorithms that are dependent on periodicity and required a normalisation for the overall amplitude differences. Similar periodicity-dependent techniques have been used by other authors (e.g. harmonics-to-noise ratio in Yumoto, Gould, & Baer, 1982; see discussion in Hanson, 1997, p.474). As is well-known, (pre-)aspiration can be periodic (Koenig, Mencl, & Lucero, 2005), but often it is not towards the offset of preaspirated parts (compare the differences in Figures 1 - 3). Therefore, in this study, we developed a *periodicity-independent* automatic measure of aspiration derived from the standard zero-crossing rate measure in the time domain, complemented by a set of acoustic measures (including H1*-H2*) corresponding to breathiness at the glottal level proposed by Hanson (1997). Additionally, we use a new voicing offset normalisation measure suitable for Vf-sequences which are variable in duration.

2. Experimental Study

2.1 Method

2.1.1 Subjects and Recordings

Data were gathered from five female (F1- F5) and five male (M1-M5) speakers of Scottish Standard English. All speakers were of Middle Class background and were between 23 and 50 years old. All SSE speakers were long term residents of Edinburgh, Nine speakers were born in the Lowland Central Belt. One speaker (F1) was born in Aberdeen.

The recordings were made in a sound-treated recording studio using an omnidirectional condenser microphone. The recording volume settings and each subject's distance from the microphone were kept constant. The subjects read a set of sentences containing target words from the computer screen. No specific instructions were provided towards the pitch accent placement in the utterances. The subject's speech rate was controlled by the prompt sentences being made to appear at regular time intervals. The preaspiration materials contained interspersed utterances from three additional experiments as distractors. The data was digitised at a sampling rate of 11025 Hz with 16-bit resolution sufficient for all acoustic analyses in this study. The male data also included parallel laryngographic recording (Laryngograph Processor ™) as a control technique for the analysis of voice offset based on speech waveforms.

2.1.2 Materials

The materials enable a comparison of word-final voiced and voiceless fricatives, produced in a variety of prosodic contexts, but were not specifically designed as a homogenous whole for the study of preaspiration, and differ for male and female subjects.

Female data

The original female data in which preaspiration was found was designed as control for a child language study (Gordeeva, 2005). The materials varied phrasal accent location and voicing, and contained three vowel heights: i.e. close, mid-close and mid-open. A subset of the complete data included six target words each repeated five times over four positions (two phrase-final and two non-final) summarised in Table 1. The carrier sentences were of the form "A fish is a fish, and nothing but a fish." or "It's a fish". This yielded 100 (5x4x5) tokens per female speaker, and a total of 500 tokens for all speakers.

Male data

The male data was recorded at a later stage and contained additional vowel heights and three minimal voicing contrast pairs (see Table 1). The vowel height data included nine target words repeated over two phrasal positions, yielding 18 (9x2) tokens per speaker, and a total of 90 tokens. The voicing contrast pairs contained six words recorded over four phrasal positions, yielding 24 (6x4) tokens per speaker, a total of 120 tokens for all speakers. These materials were of the form "That's the word bus.", "I can say bus again.". For the targets involving the fricative /voice/ contrast we used the additional carriers.

Table 1. Control conditions, materials and carrier sentences used for the female and male speakers. The uppercase words indicate phrasal accent in the carrier sentences.

Female speakers: **Vowel height targets** Close: /ʉ/ goose, /ɪ/ fish Open-mid: /ʌ/ bus **Voicing contrast targets** goo**s**e/shoe**s**/choo**s**e	**Phrasal Contexts and carrier sentences** *Final 1:* It’s a <TARGET>. *Non-final 1, Non-final 2 and Final 2:* A <TARGET> is a <TARGET>, and nothing but a <TARGET>.
Male speakers: **Vowel height targets** Close: /ɪ/ fish, dish Close-mid: /e/ Beth, place, base Open-mid: /ɛ/ best, /ɔ/ boss, /ʌ/ bus Open: /a/ bath	**Phrasal Contexts and carrier sentences** *Final 1:* That’s the word <TARGET>. *Non-final 1:* I can say <target> AGAIN.
Voicing contrast targets bu**s**/bu**zz** pla**c**e/play**s** ba**s**e/bay**s**	*Final 1:* That’s the word <TARGET>. *Non-final 2* and *Final 2:* I say <TARGET>, and not <TARGET>. *Non-final 1:* I can say <TARGET> AGAIN.

2.3 Analyses

2.3.1 Phonetic labelling

Phonemic transcription was annotated along with the segment duration for each target word. Preaspirated parts were separately time-marked. All analyses were done in PRAAT 4.3 (Boersma & Weenink, 2006). We used a combination of auditory and acoustic cues in the Vf-transitions to determine the onset of preaspiration: i.e. the auditory impression of whispery or glottal/supraglottal friction in the transition reflected in the spectrum by the presence of high-frequency noise other than that of the coda fricative, or fast and early offset of modal voicing in the waveform amplitude along with weakening of modal formant structure in the spectrum.

In Figure 2, the glottal fricative is 120 ms long, and the preaspiration onset is determined by the onset of non-[s] frication after the vowel. What matters here is the presence of glottal friction, not the loss of voicing, so both [h] and [ɦ], or any intermediate form, will count as preaspiration for the following lingual fricative. If the glottal fricative [ɦ] has uninterrupted voicing throughout the consonant (Löfqvist, Koenig, & McGowan, 1995; Koenig et al., 2005), any measure of voicing offset will be unable to determine the acoustic boundaries of preaspiration. In Figure 2 the voicing continues a third of the way into the aspirated part.

Preaspiration was time marked, however short, but if no preaspiration was found, the preaspiration marker was placed at the end of the vowel with zero duration. In voiced fricatives, the Vf boundary marker was determined by onset of target lingual friction in spectrograms irrespective of the voicing.

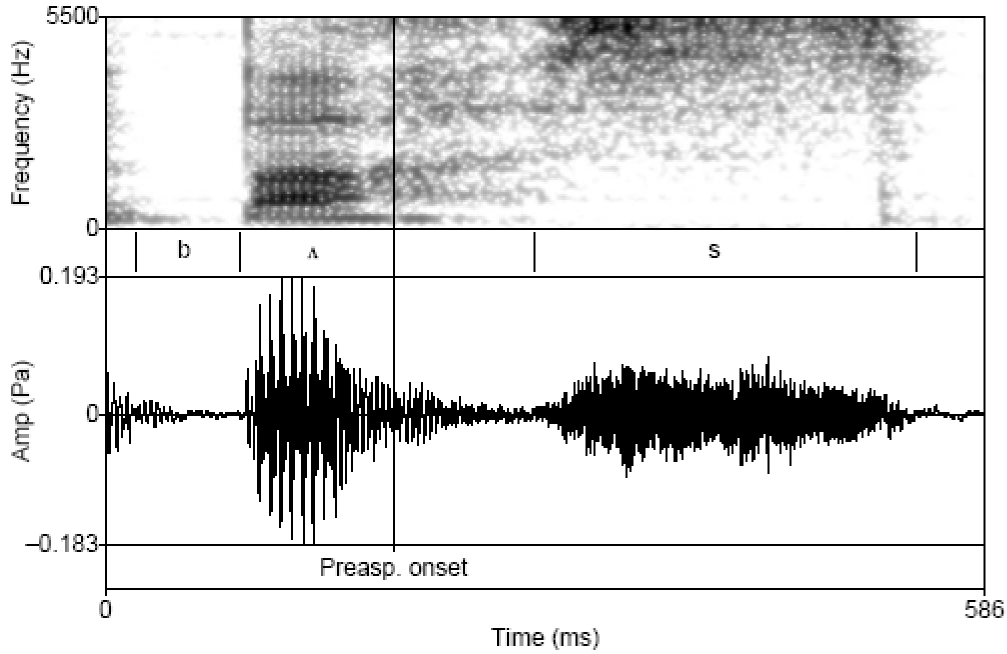


Figure 2. An example of annotation of preaspirated Vf-transition in the word “bus” (speaker F1).

In any study of Vf transitions, we would expect an early onset of laryngeal abduction relative to the supraglottal constriction. One of the initial motivations for this study was that the voicing offset in the SSE female data was so large to merit careful study. In order to obtain a categorical quantitative count of such preaspirated tokens, a threshold is required, so individual tokens were counted as preaspirated if the time delay from the preaspiration onset to the onset of the following fricative was longer than 30 ms (independent of the duration of V or f). This seems to be a reasonable durational limens to choose given following a similar perceptual threshold for preaspiration as cue to voicing contrast reported by van Dommelen (1998).

For each token we determined the presence/absence of pitch accent and accordingly labeled them as “accented” or “de-accented”, rather than relying on assumptions about how the speakers might produce the materials. Phrasal prominence was used as a selection criterion for the input in different subtests in this paper.

2.3.2 Acoustic analyses

In addition to a categorical binary analysis of preaspirated vs. non-preaspirated tokens, we performed continuous acoustic measures of preaspiration relative to Vf-boundaries in order to better understand the underlying phonetic processes. Analysis is made of the correlates of both our binary observational categorisation into preaspirated and non-preaspirated fricatives, and a binary phonological categorisation into /±voice/ tokens based solely on phonological convention.

Table 2. Overview of the acoustic measures used in this study.
Further details are given in the text.

Measure	Description
Voicing:	
voicing_offset (%)	voicing offset ratio
Duration:	
V_dur (ms)	vowel duration (including preaspiration)
f_dur (ms)	duration of the coda fricative
Preasp_dur (ms)	time from onset of preaspiration to fricative onset
Preasp_ratio	ratio Preasp_dur/V_dur
Aspiration-related measures:	
ZCR mid (per sec)	zero-crossing rate in middle (third fifth) part of the vowel
ZCR final (per sec)	ZCR in the final (fifth) part of the vowel
ZCR change (per sec)	ZCR difference between the final and middle parts of the vowel
HTN mid (dB)	harmonics-to-noise ratio in middle part of the vowel
HTN final (dB)	HTN in the final (fourth fifth) of the vowel
HTN change (dB)	HTN difference between the final and middle parts of the vowel
H1*-H2* mid (dB)	H1*-H2* ratio in the middle (third fifth) part of the vowel
H1*-H2* final (dB)	H1*-H2* in the final (fourth fifth) part of the vowel
H1*-H2* change (dB)	H1*-H2* difference between the final and middle parts of the vowel

The parameters in Table 2 were automatically derived based on manual annotations of segment duration. All acoustic analysis is performed in PRAAT 4.3. The aspiration-related parameters were measured as averages in the middle and final parts of the vowel, or as parameter change in the last part of the vowel relative to the middle vowel part, as defined in Table 2.

Voicing offset ratio (%) reflects the timing of voicing offset in the V or f part relative to the onset of final fricative (0%) on the scale between 100 and -100%, whereby the V onset marks 100% and the f-offset marks -100% (see Figure 3). This new method has the advantage of indicating a binary presence/absence of phonetic voicing (i.e. the positive and negative scaling) and its amount in percent, while simultaneously normalizing for different durations of V or f independently from each other. This method is preferable to voicing offset being measured relative to the entire Vf section taken as 100% of the rhyme, or of vowel or consonantal duration separately, as they lack the simultaneous normalisation vowel and consonantal duration relative to the timing of Vf boundary.

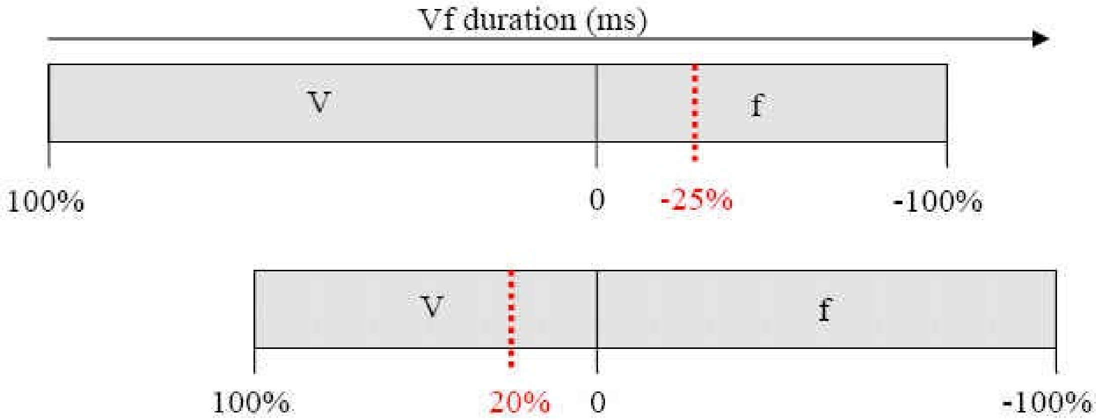


Figure 3. Visual representation of voicing offset ratio measure. The two grey thick bars represent absolute Vf duration of two different Vf-sequences. The vertical 0 markers show the oral fricative onset time, so the two grey bars are aligned at this point. The vertical red dotted lines represent the timing of voicing offset in the Vf-part. The percentages below are calculated between 0 and the fricative offset (-100%), or 0 and vowel onset (100%), depending on whether voicing offsets in the fricative (as in the upper bar which represents a partially voiced fricative) or in the vowel (as in the lower bar which represents a preaspirated voiceless fricative).

Periodicity was measured from speech waveform using the cross-correlation algorithm with the minimum of 75 Hz and maximum of 350 Hz for the male data, and 75 Hz and 400 Hz for the female data. The minima and maxima were based on F0 vowel ranges in both groups. Prior to this, speech waveforms were high-pass filtered at 50 Hz to get rid of the DC component. The performance of the input parameters for periodicity measure from speech waveforms was further calibrated against periodicity measured from the available subset of EGG waveform to achieve the best correlations (see reliability for further discussion).

Zero-crossing rate (ZCR per sec) as implemented here reflects the amount of aspiration/breathiness in the spectrum above the fundamental. ZCR is a standard measure calculated in the time-domain of a waveform as the number of zero-crossings of the wave within a certain part of signal, divided by the number of samples in this part (e.g. Rowden, 1992, p.45-46). ZCR tends to be the highest for voiceless fricatives. Waveforms were band-pass filtered with an upper limit at 5.5 kHz (i.e. Nyquist frequency) and a flexible lower limit (defined at 1.5*maximum pitch for each vowel token) designed to remove low frequency deviations away from the zero-line due to potential presence of voicing. In breathy speech, such low frequency (H1) components have the effect of quasi-sinusoidal displacing the wave away from the zero-line (Klatt et al., 1990; Stevens, 1998) (see Pane A in Figure 4). Without high-pass filtering the presence of low harmonics renders a ZCR measure meaningless as an indicator of midrange aperiodic noise. The effect of high-pass filtering is illustrated in Figure 4 on Pane B, where the quasi-sinusoidal H1 domination in the time-domain is reduced, and the zero-crossings represent the midrange non-modal breathiness visible in the upper spectrogram pane. Zero-crossings are counted in frames of 10 ms.

ZCR should be a better indicator of the midrange noise than more traditional spectral measures such as spectral tilt (H1*-H3*) (Hanson, 1997), as the H3 component (dB) contains spectral level contributions from both periodic and aperiodic components, while higher ZCR from high-pass filtered waveform in the time-domain mainly reflects noise. This measure of breathiness/aspiration has further advantage of being independent of the accuracy of pitch trackers, it does not require amplitude normalisation (as e.g. in Hillenbrand et al., 1994) and is fully automatic and is easy to compute.

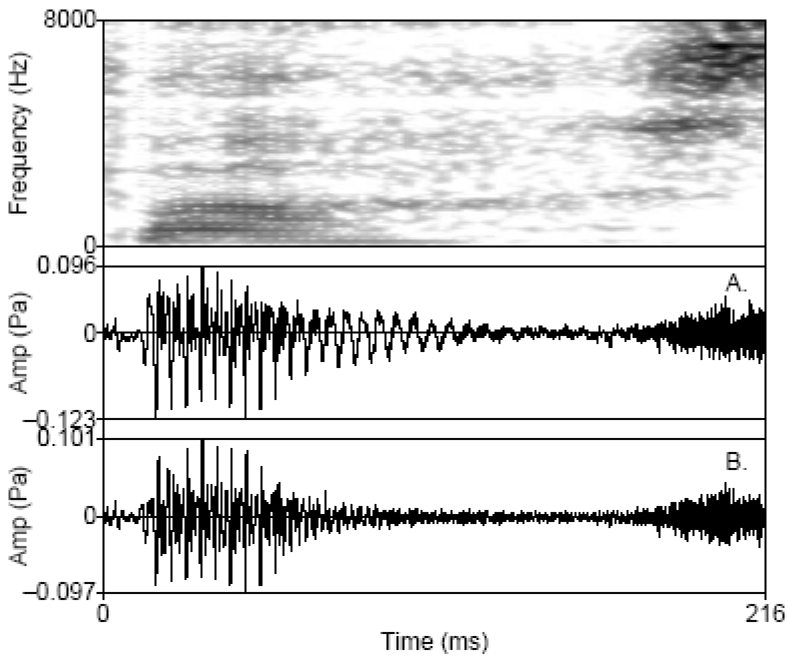


Figure 4. The effect of high-pass filtering on zero-crossing rate in a time domain waveform with breathy vowel /ʌ/ followed by fricative /s/: the unfiltered waveform with sinusoidal H1domination in the Vf-transition is represented in Pane A., the band-pass filtered form in Pane B, the upper pane represents the spectrogram of the unfiltered speech.

Harmonics-to-noise ratio (HNR, dB) reflects a relative amount of aperiodicity in otherwise periodic portions of speech spectrum. This ratio does not show which aperiodic component (noise, jitter or shimmer) contributes to aperiodicity (Murphy, 1999), and as such can be a correlate of hoarseness (Yumoto et al., 1982) as well as aspiration. An HNR ratio of 0 dB means that there is equal energy in the harmonics and noise, while 1% of noise equals 20 dB. HNR is calculated here using the Harmonicity autocorrelation method in PRAAT in frames of 10 ms and minimal pitch of 65 Hz. We only performed HNR comparisons in the first 80% of the vowel, since the final 20% often lacked periodicity due to voiceless preaspiration, in which case the ratio is incomputable.

H1*-H2* (dB) is the difference between the levels of the first (H1) and second (H2) harmonics of the vowel. It serves as an indication of open quotient (OQ): i.e. the timing of open phase of the glottal cycle relative to the total time of the period. Due to the increase of the glottal area, larger OQ should lead to greater losses and more aspiration noise (Holmberg et al. 1995). The more breathy types of phonation are characterised by lower values of H2 relative to H1 (Hanson, 1997; Löfqvist, 1995). H1 and H2 levels were measured from the spectrum in a Hamming window with a window length covering two pitch periods. Raw H1 and H2 values were then corrected for the ‘boost’ effect resulting the proximity of the first formant following the procedure in Hanson (1997, p.475). This resulted in boost-effect corrected H1*-H2* values.

2.3.3 Statistical analyses

Distributional characteristics (vowel height, phrasal position, sex) are assessed by means of non-parametric Chi-square (χ^2) tests at .05 level of confidence for the subsets of data fulfilling the validity requirements of 'expected frequencies'. To establish the hierarchy of acoustic correlates of preaspiration and final fricative /voice/, we followed the following procedure.

Multivariate Analysis of Variance (MANOVA, α = .05) was used with acoustic variables in Table 2 as dependent variables, and with the categories PREASPIRATION or VOICE as fixed factors to determine which of the variables have a significant effect on either of the fixed factors. Subsequently Linear Discriminant Analysis (LDA) was used to determine the relative ranking of each of the significant variables in predicting PREASPIRATION or VOICE. The 'stepwise' LDA was chosen, since it makes no assumptions about which predictor should have higher priority than others, and the order of predictor entry is determined by statistical criteria (Wilks' Lambda with an F value of 3.84 for predictor entry and 2.71 for removal).

2.3.4 Reliability

In this study, the timing of voicing offset in the Vf transition is measured from speech waveforms by the first author, using the landmarks and procedures described above. Instrumental techniques like electroglottography (EGG) are known to more reliably determine voicing at the voice source (Marasek, 1997; Smith, 1997). The reliability of the voicing offset measure in the speech waveform was tested against the EGG data simultaneously acquired for the male speakers with the Laryngograph Processor. Using the same input procedure as for speech waveforms, we derived the voicing offset timing from the EGG signal, with the difference that EGG waveforms were digitised at 8000 Hz (16-bit resolution), and were pass-band filtered (Hann method) between 60 and 4000 Hz.

We performed a consistency check of voicing offset timing and voicing offset ratio measured from EGG and the corresponding speech waveforms. The root-mean-square (RMS) error between the voicing offset timing in EGG and acoustic waveform was 16.4 ms, and corresponded to the mean 4% of the Vf duration. Additionally, the voicing offset ratio based on EGG and acoustic waveforms had a highly significant correlation [$r=0.88$; $n=203$; $p<.0001$], proving a good agreement between the two measures.

In order to evaluate the consistency of the manual timing annotations, the first author re-measured the timing of the Vf and preaspiration boundaries eight months after the original analysis. The test was based on a random 10% of the data, both male and female. The RMS-error for the Vf-boundary and onset of preaspiration marker was 8.3 ms and 8 ms accordingly, and corresponded to only a small mean 1.6% and 1.5% of the total Vf- duration. Based on these results, we considered the manual annotation of segment duration in the acoustic data from both males and females to be reliable.

2.4 Results

2.4.1 Summary

Non-parametric results show that in terms of frequency of occurrence, preaspiration in SSE word-final Vf-transitions is not a sex-related, but rather a speaker-dependent characteristic, although female speakers have significantly longer Preasp_ratio than male speakers. Preaspiration occurs more frequently (although not exclusively) in more open vowels and in phrase-final positions.

In acoustic terms, preaspiration occurs in vowels with significantly longer duration than in non-preaspirated variants. Furthermore, it is acoustically shaped by midrange spectral aperiodicity expressed as zero-crossing rate in the final part of the vowel or as more abrupt ZCR changes in the second half of the vowel.

The hypothesis that preaspiration functions as an important correlate of fricative /voice/ is also supported in this study. In fact, the most successful predictor of /voice/ in phrase-final fricatives is the rapid increase of zero-crossing rate (i.e. higher frequency noise) in the second half of the vowel, and its high magnitude in Vf-transitions. The importance of zero-crossing rate in cueing the fricative /voice/ contrast surpasses in strength the traditionally considered parameters such as voicing offset and consonantal and vowel duration.

The following sections describe these findings in more detail.

2.4.2 Phonetic conditioning of preaspiration

2.4.2.1 Speaker's sex differences or individual variation?

Using the categorical definition of preaspiration, we compared the distributions of preaspirated and non-preaspirated realisations of all target words (*n*=438) with voiceless fricatives produced by female and male speakers. No significant sex differences [χ^2=.5; df=1; p=.48] are found in the distributions: the descriptive statistics can be found in Table 3. The female speakers preaspirated 41 % of their fricatives and the males 45%. Individual speaker variation in the percentage of occurrence is shown in Figure 5.

Figure 5 shows that all speakers, male or female, produced preaspirated variants – with the speakers F1 and M5 producing preaspirated fricatives almost exclusively.

Table 3. Distributions (number of tokens and percentages) of preaspirated and non-preaspirated variants by speakers' sex.

			Preaspiration		Total
			No	Yes	
Sex	Female	Count	176	124	300
		% within sex	59	41	100
		% of Total	40	28	68
	Male	Count	76	62	138
		% within sex	55	45	100
		% of Total	17	14	32
Total		Count	252	186	438
		% within sex	58	42	100
		% of Total	58	42	100

Although there are no significant sex differences in terms of frequency of occurrence, a separate t-test shows that the female speakers produce significantly longer Preasp_ratio compared to male subjects [$F=83.5$; $df=1$; $p<.001$] across all preaspirated voiceless fricatives. The means and standard deviations of sex-specific preaspiration ratios are reported in Table 4. Since Preasp_ratio includes vowel and preaspiration, the ratio of 0.5 means that preaspiration is as long as the full vowel. Individual female speaker means range between 0.33 and 0.46 approaching the 0.5 value. This result confirms the trend widely reported in the literature (e.g. Henton & Bladon, 1985; Fant, Kruckenberg, & Nord, 1991; Hanson, 1997) that females have more breathy phonation than male speakers.

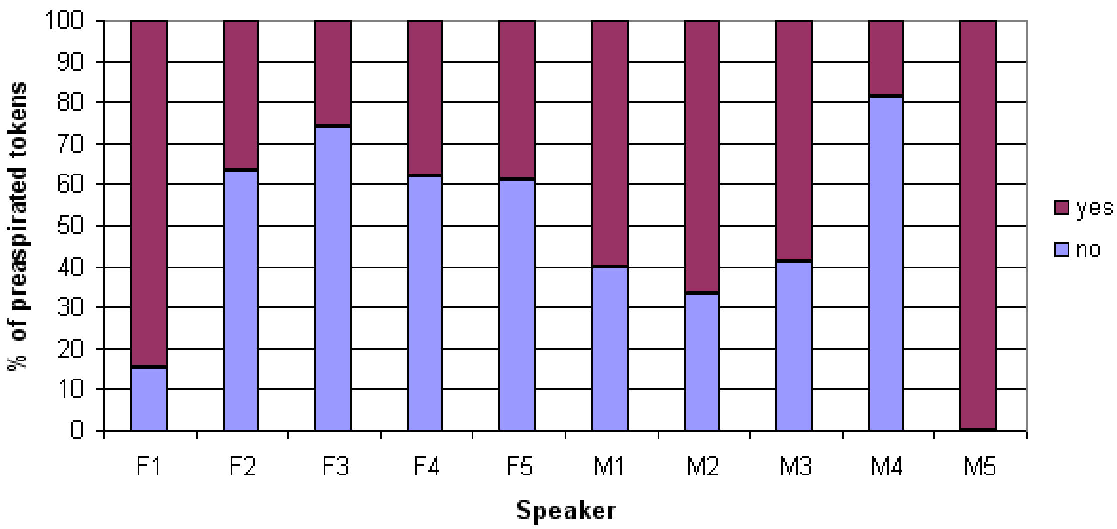


Figure 5. Percentage of preaspirated ("yes") and non-preaspirated Vf-tokens per speaker.

Table 4. Descriptive statistics for Preasp_ratio across all preaspirated voiceless fricatives for the individual speakers and by speakers' sex.

Sex	Speaker	Mean	Std. Deviation	N
Female	F1	0.45	0.09	48
	F2	0.34	0.09	21
	F3	0.35	0.08	17
	F4	0.46	0.09	16
	F5	0.33	0.08	22
	Total	0.40	0.10	124
Male	M1	0.26	0.05	11
	M2	0.26	0.04	12
	M3	0.27	0.05	14
	M4	0.26	0.08	7
	M5	0.28	0.05	18
	Total	0.27	0.05	62
Total		0.35	0.11	186

Overall the ranges of Preasp_ratio are somewhat lower than the similarly calculated ratios for preaspirated fricatives reported in Jones and Llamas (2003) ranging from 0.51 to 0.62 in three Middlesbrough English speakers. This is not surprising given that Middlesbrough English is a generally more preaspirating variety (e.g. also in stops), than SSE.

2.4.2.2 Effect of Vowel Height

The test of association between phonological vowel height (close, mid-close, mid-open, open) and our categorical analysis of the presence of preaspiration is performed on the subset of data containing coda /-voice/ fricatives. Separate tests were run for the male and female data, because of the differences in vowel heights available in the subsets. The results indicate that there is a highly significant association between the occurrence of preaspirated fricatives and vowel height in both male [χ^2=27.308; df=3; p <0.001] and female data [χ^2=64.03; df=1; p <0.001]. The distributions of preaspirated and non-preaspirated fricatives are shown per vowel in Figure 6 across all speakers. Table 5 shows that the number of preaspirated variants increases with a more open vowel height in both male and female speakers.

Table 5. Distributions of preaspirated and non-preaspirated voiceless fricatives as a function of vowel height in male and female speakers.

			Preaspiration		
			no	yes	Total
Male speakers					
Vowel height	close	Count	23	2	25
		% within V height	92	8	100
	mid-close	Count	40	28	68
		% within V height	59	41	100
	mid-open	Count	11	23	34
		% within V height	32	68	100
	open	Count	2	9	11
		% within V height	18	82	100
Female speakers					
Vowel height	close	Count	149	50	199
		% within V height	75	25	100
	mid-open	Count	27	74	101
		% within V height	27	73	100

This vowel height dependence seems to (at least partly) be explained by the tendency of open vowels to have longer intrinsic duration, as there is also a significant positive correlation between duration of preaspiration and the total vowel duration [$r=0.62$; $n=386$; $p<.0001$].

Jones and Llamas (2003) for Middlesbrough English reported the mean proportion of preaspiration to the total vowel duration in “mass” of 0.54. The proportion is similar before /-voice/ stops and fricatives. The mean Preasp_ratio in our data is a somewhat lower 0.31 for open and mid-open vowels across all phrasal positions.

Helgason (2002) does not report any statistics for preaspirated fricatives for Central Swedish. However, the Preasp_ratio before voiceless stops for read speech is very similar to SSE, and ranges from 0.32 to 0.37 in phase-accented positions.

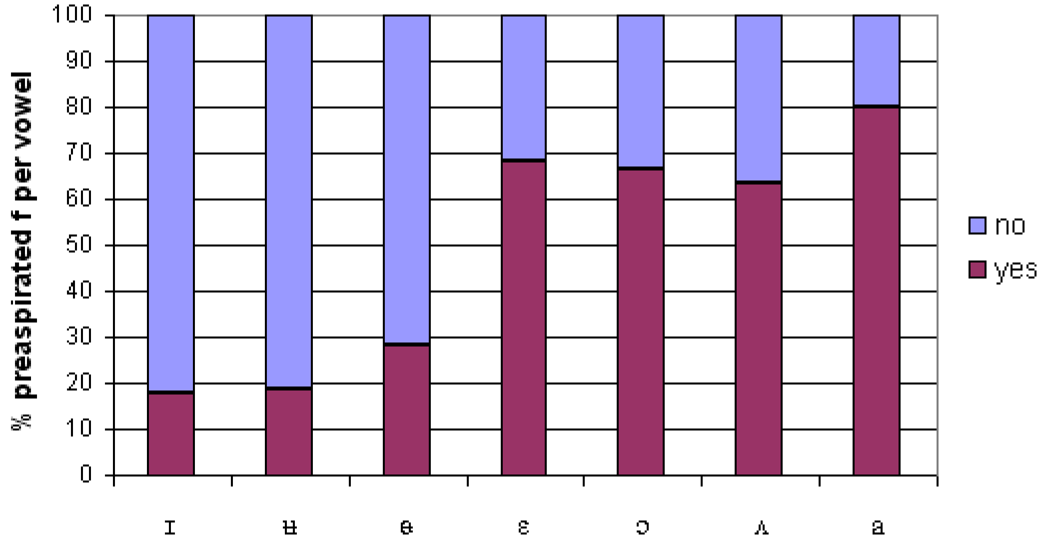


Figure 6. Percentages of preaspirated and non-preaspirated fricatives per vowel across all speakers.

2.4.2.3 Phrasal Position

The test of the effect of position in phrase (non-final and final) on the occurrence of preaspirated variants is measured, just for the subset of words (*n*=342) carrying phrasal accent. The results (see Table 6) show a highly significant association [χ^2=12.514; df=1; $p < 0.001$] between the occurrence of preaspirated fricatives and the phrasal position of the target, with an overall 20% higher rate of preaspirated variants appearing in phrase-final accented positions compared to non-final ones. This result holds for both sexes (see Figure 7). The distributions of four tested phrasal positions (see Table 1 for content) per speakers' sex are shown in Table 7.

Table 6. Distributions of preaspirated and non-preaspirated fricatives as a function of phrasal position across all speakers.

			Phrasal Position (PP)		Total
			non-final	Final	
Preaspirated?	No	Count	105	80	185
		% of PP	64	45	54
		% of Total	31	23	54
	Yes	Count	59	98	157
		% of PP	36	55	46
		% of Total	17	29	46
Total		Count	164	178	342
		% of PP	100	100	100
		% of Total	48	52	100

Table 7. Distributions of preaspirated and non-preaspirated fricatives as a function of phrasal position across per speaker sex.

Male Speakers			Phrasal Position (PP) of Vf tokens				Total
			non-final 1	non-final 2	final 1	final 2	
Preaspirated?	No	N tokens	24	8	13	4	49
		% of PP	100	50	29	27	49
	Yes	N tokens	0	8	32	11	51
		% of PP	0	50	71	73	51
	Total	N tokens	24	16	45	15	100
		% of PP	100	100	100	100	100
		% Total	24	16	45	15	100
Female Speakers			non-final 1	non-final 2	final 1	final 2	
Preaspirated?	No	N tokens	42	31	33	30	136
		% of PP	64	53	44	70	56
	Yes	N tokens	24	27	42	13	106
		% of PP	36	47	56	30	44
	Total	N tokens	66	58	75	43	242
		% of PP	100	100	100	100	100
		% Total	27	24	31	18	100

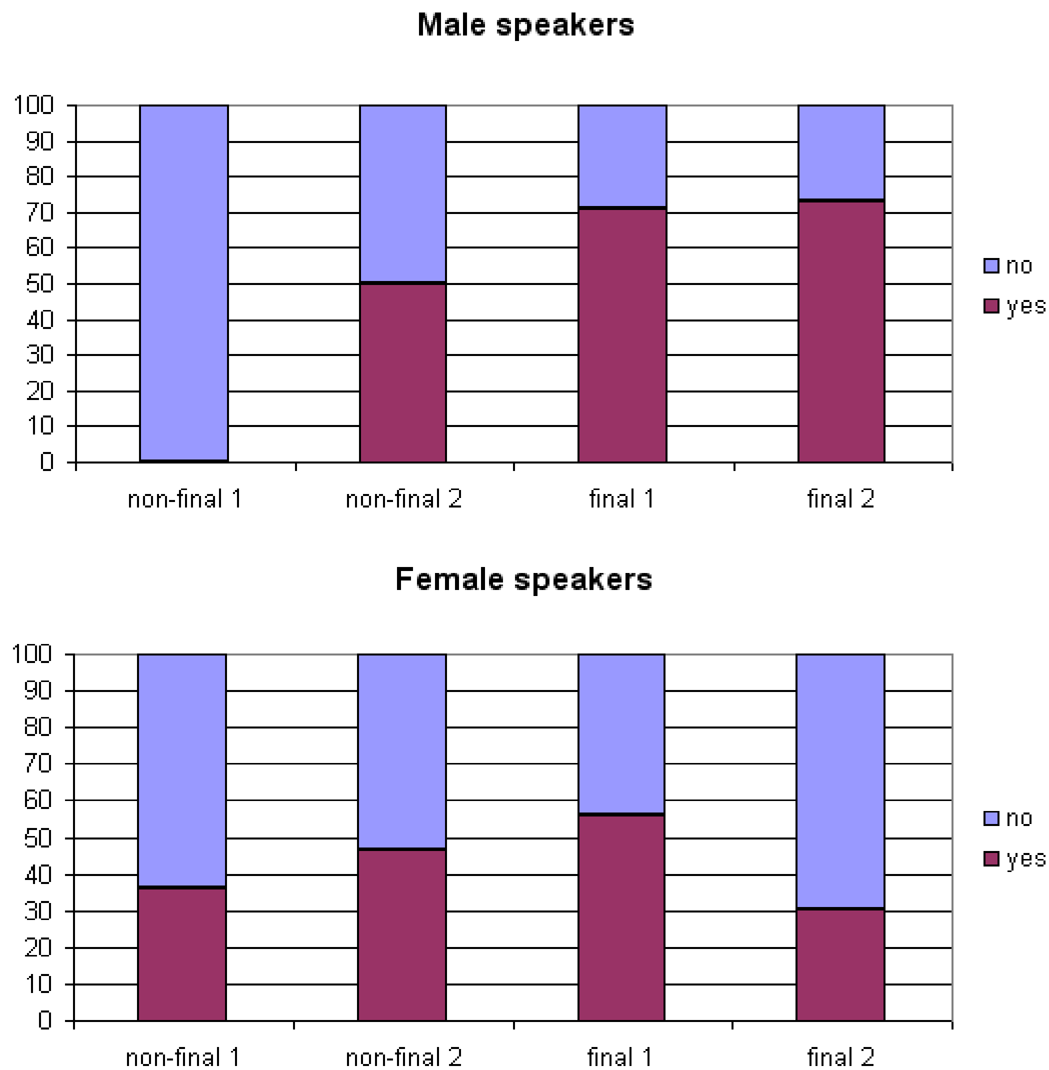


Figure 7. Percentages of preaspirated and non-preaspirated fricatives per phrasal position for female and male speakers. Note that the content of materials for the four labelled positions is somewhat different between the sexes.

Figure 7 shows that the male speakers tended to have bigger gaps than females in the frequency distributions of preaspiration between the phrase non-final and final contexts. This sex difference is partly explainable by the fact that female speakers produced more preaspirated variants in non-final positions. Besides, female speakers produced fewer preaspirated transitions in the Final 2 context, which could be partly explained by the fact that this phrase-final context (unlike Final 2 in the male materials) involved much longer utterances than Final 1.

The general results both across and per speaker sex show that phrase-final positions trigger significantly more occurrences of preaspirated variants than the non-final ones.

Table 7. Distributions of preaspirated and non-preaspirated fricatives as a function of phrasal position across per speaker sex.

Male Speakers			Phrasal Position (PP) of Vf tokens				Total
			non-final 1	non-final 2	final 1	final 2	
Preaspirated?	No	N tokens	24	8	13	4	49
		% of PP	100	50	29	27	49
	Yes	N tokens	0	8	32	11	51
		% of PP	0	50	71	73	51
	Total	N tokens	24	16	45	15	100
		% of PP	100	100	100	100	100
		% Total	24	16	45	15	100
Female Speakers			non-final 1	non-final 2	final 1	final 2	
Preaspirated?	No	N tokens	42	31	33	30	136
		% of PP	64	53	44	70	56
	Yes	N tokens	24	27	42	13	106
		% of PP	36	47	56	30	44
	Total	N tokens	66	58	75	43	242
		% of PP	100	100	100	100	100
		% Total	27	24	31	18	100

2.4.3 Hierarchy of the acoustic correlates of preaspirated fricatives

In this part, we would like to approach this phenomenon from a different angle by establishing the hierarchy of acoustic correlates of preaspiration found in SSE vowel - voiceless fricatives transitions.

We run Multivariate Analysis of Variance (MANOVA) with the acoustic measures (other than Preasp_ratio) in Table 2 as independent variables and PREASPIRATION (“yes” or “no”) as a fixed factor to determine which of the acoustic variables are significantly affected by preaspiration. The analysis was carried out on a selection of all instances carrying phrasal accent and ending with voiceless fricatives (n=298). Table 8 summarises the acoustic variables that are significantly affected by preaspiration with means and standard deviations, as well as the significance levels (with F- and p-values) for those variables.

The consonantal duration and H1*-H2* in the mid part of the vowel were not significantly affected.

In order to measure the ability of each of the significant acoustic correlates to predict preaspiration and establish their ranking, we subjected the acoustic measures to stepwise Linear Discriminant Analysis (LDA). The same targets as for MANOVA (n=298) were entered in SPSS 12 Discriminant Analysis. PREASPIRATION (“yes” or “no”) was used as the predicted variable, and the acoustic measures in Table 8 as independent variables (predictors).

Table 8. Means, one standard deviation, and MANOVA results for the acoustic variables significantly affected by preaspiration pooled across all subjects.

Acoustic variables	Preaspirated? (Total *n*=298)				Significance	
	No (*n*=178)		Yes (*n*=120)		df=1	
	Mean	Stddev	Mean	Stddev	F	p
V_dur (ms)	123	30	170	46	114.9	0.000
Voicing offset ratio (%)	-9.2	18.8	2.7	14.5	34.0	0.000
ZCR mid (per sec)	1119	329	1488	580	49.0	0.000
ZCR final (per sec)	2271	1131	3444	1201	73.4	0.000
ZCR change (per sec)	1152	1050	1956	1119	39.9	0.000
HTN mid (dB)	15.2	5.2	12.7	4.2	18.7	0.000
HTN final (dB)	13.6	4.4	10.1	4.7	41.8	0.000
HTN change (dB)	1.6	9.6	3.9	7.4	4.9	0.028
H1*-H2* final (dB)	10.0	8.6	12.1	8.3	4.0	0.046
H1*-H2* change (dB)	1.6	9.6	3.9	7.4	4.9	0.028

The results indicate that 84.9 % of all targets were correctly classified into our annotation-based categories of non- or preaspirated based on the acoustic measures. The percentage is well above the chance level (70%). The relative ranking of the acoustic variables is summarised in Table 9, where they are ordered by their correlation size with standardised canonical discriminant functions. The best predictors of the presence or absence of PREASPIRATION are vowel duration, zero-crossing rate in the final vowel part and its change in the second part of the vowel, followed by the voicing offset and harmonics-to-noise ratio in the second part of the vowel.

Individual speaker means for preaspirated and non-preaspirated Vf transitions for the 5 selected LDA parameters and HTN final measure are plotted in Figure 8. To give an idea about the individual and sex differences for each of the parameters for the means of individual female speakers are plotted in solid lines, while the male speakers are represented by dotted lines.

The figure shows that preaspirated variants have longer duration; have substantially higher ZCR values (vowel-finally and as change in the second part of the vowel) reflecting the increasing aspiration noise levels above the fundamental frequency; have less periodicity compared to noise (lower HTN values), and that the voicing offsets earlier relative to Vf-boundaries. Although there are both individual and sex differences for most of the parameters, the individual speakers consistently produce the same direction of the differences for the top four LDA parameters reflecting coherency of the MANOVA and LDA results. There is more individual variation in HTN final measure (with the speaker F4 producing similar levels for non- and preaspirated variants), and H1*-H2* change measure.

Table 9. Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions for PREASPIRATION. Variables are ordered by absolute size of correlation within function. Variables marked with (a) were not selected by LDA as predictors.

Acoustic variable	
V_dur	0.627
ZCR final	0.501
ZCR mid (a)	0.409
ZCR change	0.369
Voicing offset ratio	0.341
HTN final (a)	-0.268
HTN mid (a)	-0.261
H1*-H2* change (a)	0.129
HTN change	0.129
H1*-H2* final (a)	0.012

Speaker M5 produced only preaspirated realisations under phrasal accent, therefore this subject's data misses the non-preaspirated data points.

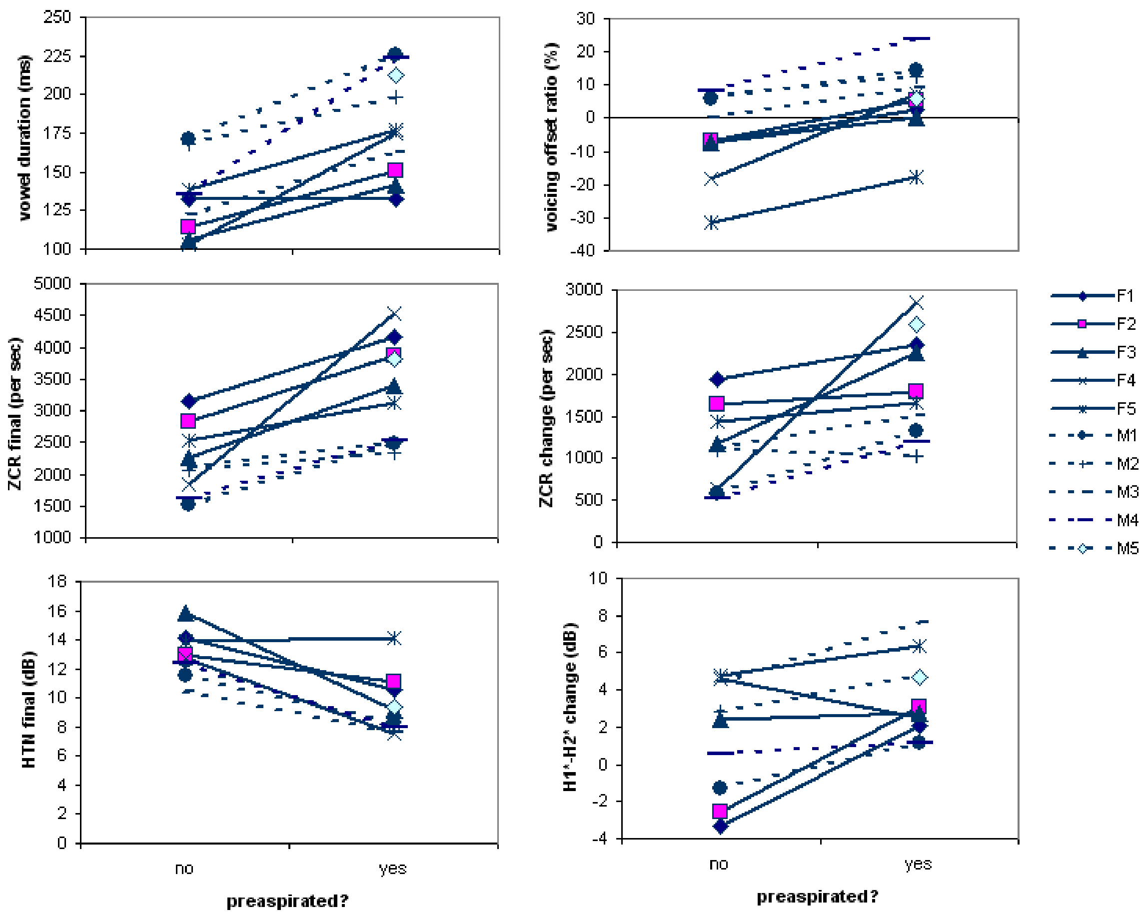


Figure 8. Individual speaker means of preaspirated and non-preaspirated voiceless fricatives for six acoustic predictors of preaspiration.

Figure 8 also shows some sex-related differences between the individual speakers. For the two ZCR measures the female subjects produce higher ZCR values (more high frequency aspiration) consistently with the literature on sex-specific voice quality (Klatt et al., 1990; Fant et al., 1991; Hillenbrand et al., 1994; Hanson, 1997). The sex differences in vowel duration reflect differences in materials (female data contained more close vowels). The higher HTN values (reflecting more periodicity than males) can be a result of creakier phonation type in male subjects, but can not be interpreted with certainty because of the already discussed ambiguity of the HTN measure as to breathiness or creakiness (Yumoto et al., 1982). The sex differences in voicing offset ratio could be an artefact of voicing offset measure being calibrated against male EGG data, and can, therefore, not be interpreted as a relevant sex-related difference (although the *relative* differences in voicing offset within each individual subject keep high coherence).

The fact that longer vowel duration is associated with the occurrence of preaspiration in Vf-transitions is not surprising given the non-parametric results above on the greater likelihood of preaspiration to occur after more open vowels (i.e. intrinsically longer), and its bigger chance of appearing in phrase-final positions, a context which triggers longer segmental durations marking prosodic edges (Edwards et al., 1991; Cho, 2001). So the primacy of vowel duration in cueing preaspiration is likely to reflect these phonetic/prosodic distributions.

Apart from that, preaspiration in the SSE Vf-transitions is best predicted by aspiration-related parameters in the final part of the vowel and its change throughout the second half of the vowel, voicing offset ratio and harmonics-to-noise ratio change. The fact that ZCR substantially increases in the preaspirated cases towards the end of the vowel, while the HTN-ratio decreases, reflects a substantially smaller amount of periodicity compared to noise in the second half of the preaspirated vowel. The individual speakers in Figure 8 are consistent in realising the directions of differences irrespective of subject and sex.

Although the H1*-H2* final and change measures were significantly affected by preaspiration in having about 2 dB bigger differences between H1 and H2 (suggesting breathier offset due to somewhat bigger open quotient), the measures did not contribute sufficiently to the % correct classification and were amongst the lowest in ranking. This fact supports previous reports in the literature that, in so far H1*-H2* reflects glottal opening, it seems to be a parameter independent from the parameters reflecting the *amount* of glottal airflow (Klatt et al., 1990; Hanson, 1997) reflected in the zero crossing rate (ZCR) in this study. It suggests that it is possible to adjust OQ without increases in the glottal airflow, as well as adjust the amount of glottal flow without changes in OQ.

Overall, the results show that preaspiration is primarily shaped by the acoustic parameters reflecting aspiration in higher spectral frequencies.

2.4.4 Hierarchy of the acoustic correlates of phrase-final fricative /voice/

In this part, we explore the hypothesis that preaspiration functions to enhance phonological /voice/ contrast in phrase-final CVf fricatives.

We run Multivariate Analysis of Variance (MANOVA) with acoustic measures (apart from Preasp_ratio) in Table 2 as independent variables, and VOICE ("yes" or "no") as a fixed factor to determine which of the acoustic variables are significantly affected by underlying fricative /voice/. The analysis was carried out on targets in phrase-final positions realised with a phrasal accent and ending with fricatives (*n*=358). Table 10 summarises the acoustic variables that are significantly affected by VOICE with means and standard deviations for the affected acoustic measures in the Vf sequences, as well as the significance levels (with F- and p-values) for those variables.

Table 10. Means, one standard deviation, and MANOVA results for the acoustic variables significantly affected by VOICE pooled across all subjects.

	Voice (Total *n*=358)				Significance	
	No (*n*=147)		Yes (*n*=211)			df=1
Acoustic variable	Mean	Stddev	Mean	Stddev	F	P
V_dur (ms)	154	51	200	57	62.8	0.000
F_dur (ms)	207	66	165	61	38.2	0.000
Voicing offset ratio (%)	-2.4	18.4	-28.4	21.4	142.4	0.000
ZCR final (per sec)	3064	1215	1545	953	175.1	0.000
ZCR change (per sec)	1727	1144	206	944	188.6	0.000
HTN mid (dB)	14.4	5.3	18.1	5.6	38.3	0.000
HTN final (dB)	11.8	5.0	17.7	5.8	101.9	0.000
H1*-H2* final (dB)	10.0	8.0	10.7	10.8	3.8	0.053

The measures ZCR mid, H1*-H2* mid and change and HTN change were not significantly affected by VOICE.

In order to measure the ability of each of the significant acoustic correlates to predict phonological /voice/ and establish their ranking, we subjected the acoustic measures to stepwise LDA. The same targets as for MANOVA (*n*=358) were entered in SPSS 12 Discriminant Analysis. VOICE ("yes" or "no") was used as predicted variable, and the acoustic measures in Table 10 as independent variables (predictors).

The results indicate that 92.2 % of all phonologically voiced or voiceless fricatives were correctly classified. The % correct classification is well above the chance level (70%). The ranking of the acoustic measures is listed in Table 11.

The results in Table 11 show that zero-crossing rate change through the second half of vowel and its amount at the end of the vowel are the most successful predictors of the phrase-final fricative voicing, followed by the voicing offset ratio. The importance of voicing offset, and vowel and consonantal duration in cueing fricative /voice/ is well known from the literature (Haggard, 1978; Docherty, 1992; Smith, 1997), and is also corroborated in this study since all these parameters are significantly affected by fricative VOICE in the MANOVA. However, the higher importance of the transitional aspiration-related parameter ZCR in cueing phrase-final fricative /voice/ has so far not been attested.

Table 11. Pooled within-groups correlations between discriminating variables and standardised canonical discriminant functions for VOICE. Variables are ordered by absolute size of correlation within function. Variables marked with (a) were not selected by LDA as predictors contributing to % correct classification.

Acoustic variable	Correlations
ZCR change	-0.518
ZCR final	-0.500
Voicing offset ratio	-0.451
HTN final	0.381
V_dur	0.299
HTN mid	0.234
F_dur (a)	-0.148
H1*-H2* final (a)	-0.070

The result supports our hypothesis that given the fact that phonetic devoicing of voiced fricatives is likely to occur in phrase-final positions, it is also likely that transitional aspiration helps to maintain the /voice/ contrast in this prosodic context. This conclusion is also supported by the parallel significance and good LDA ranking of HTN ratio in the final part of the vowel. The durational correlates seem to score less successfully than the ZCR measures and voicing offset ratio in the % correct classification of fricative /voice/.

It is interesting to note that while the high % correct classification nearly approaches one-to-one mapping of these multidimensional phonetic correlates and underlying /voice/, the correlation strength of individual predictors in the ranking is relatively low (despite statistical significance of individual correlates). This shows that underlying /voice/ contrast is controlled by a multitude of overlapping phonetic correlates with no one-to-one correspondences, and that the correlates can dynamically gain/lose strength in certain phonetic contexts.

Individual subject's means for the top four LDA VOICE predictors and vowel and consonantal duration are plotted in Figure 9.

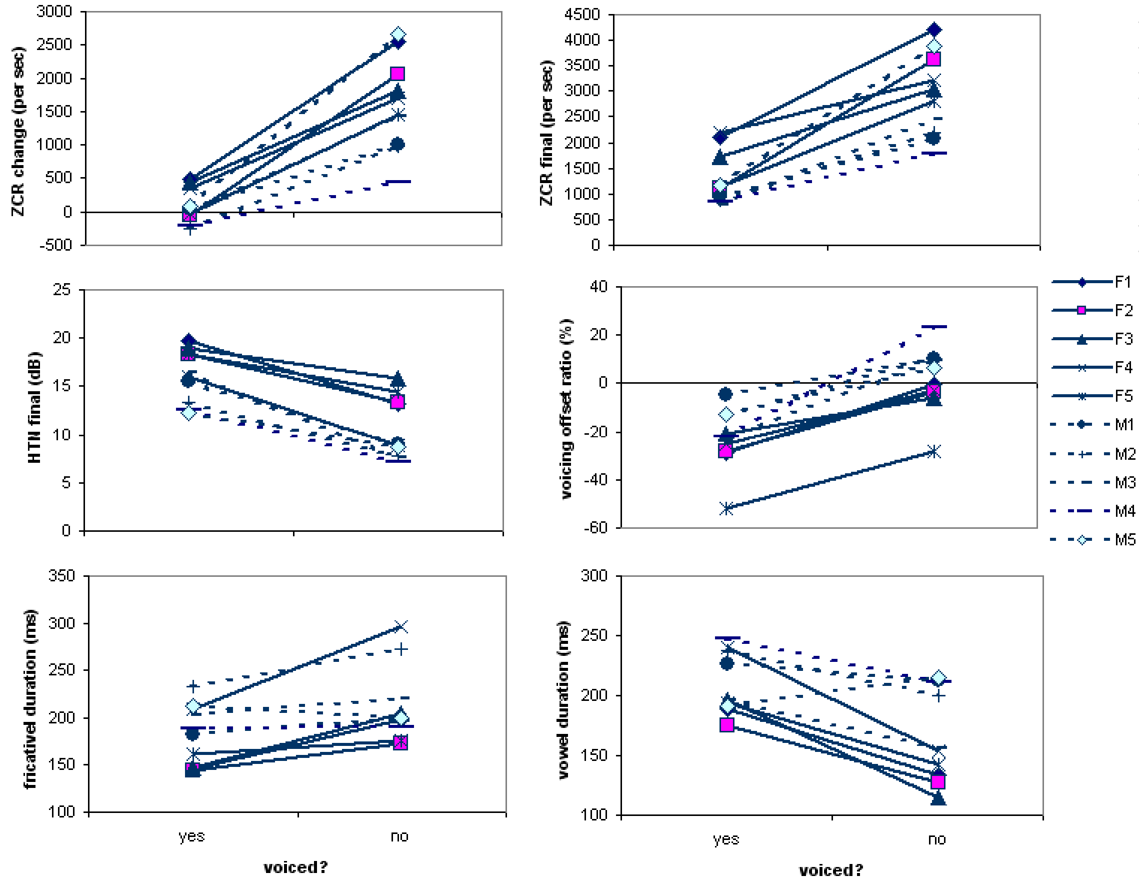


Figure 9. Individual subject means of VOICE ("yes" or "no") plotted for the four highest LDA ranked acoustic parameters and additionally vowel and consonantal duration means. Solid lines indicate female speakers. Dotted lines indicate male speakers.

The relative differences in the top four predictors of VOICE are very consistent for all individual subjects (and irrespective of sex). The direction of the differences is similar to that for PREASPIRATION: higher ZCR values in the Vf transition and in the second half of the vowel, earlier voicing offset and lower HTN values (less periodicity). The sex-related differences are also similar to those described in Figure 8.

Additionally, there are also remarkable individual differences observable in Figure 9. Male subject M5 (who produced exclusively preaspirated realisations in all tokens ending with voiceless fricatives shown in Figures 5 and 8) produces the highest difference in ZCR parameters, but unlike other subjects fails to produce the durational differences in VOICE for both vowels and consonants, suggesting that the /voice/ contrast is primarily encoded by aspiration-related parameters and less so by duration or voicing offset.

3. Discussion

Overall, this study shows that preaspiration systematically accompanies /-voice/ fricative single segment codas in SSE, and that its extent is sufficiently large (at times as long as the vowel itself) to merit the use of the phonetic label "preaspiration". Similar conclusions regarding the phonetic status of such transitions have been drawn by

Helgason for Central Standard Swedish (2002) and Jones and Llamas for Middlesbrough variety of British English (2003). However, it is noteworthy that the system of preaspiration in SSE is asymmetrical in lacking preaspiration before /-voice/ stops, where pre-glottalisation seems to be preferred. This shows that preaspiration of voiceless fricatives is not a characteristic that is per definition coupled to preaspiration in stops in a variety. It must be considered separately, at the level of co-articulation, but can be seen at a more abstract phonological level as a parallel mechanism for expressing "voicelessness".

There is still speaker variation, and contextual variation, so we would not wish to claim that this is a normative phenomenon, but preaspiration of voiceless fricatives is nevertheless important for understanding the underlying sound system of the variety. Moreover, general theoretical and methodological issues have been illuminated through consideration of this data.

3.1 Interspeaker differences and speaker's sex

Aspirated Vf transitions of longer than 30 ms equally occur in the speech of all SSE speakers in terms of frequency of occurrence irrespective of speaker's sex. However, female speakers exhibit significantly longer duration of preaspiration (in terms of Preasp_ratio) compared to male speakers. This difference is consistent with previous reports that females tend to have breathier voice quality than male speakers both in terms of laryngeal physiology (Klatt et al., 1990; Fant et al., 1991; Hillenbrand et al., 1994; Hanson, 1997) and possibly in term of specific tendency of British English females to use breathiness as a sociolinguistic marker in normal speech (Henton et al., 1985). Such Vf-transitions can be another (perhaps short-term) phonetic landmark, where such sex differences, physiological or sociolinguistic can be manifested.

Two speakers out of ten (F1: female and M5: male) produced aspirated Vf-transitions almost exclusively in terms of frequency of occurrence, with the female speaker F1 often producing preaspiration nearly as long as the vowel itself (mean Preasp_ratio of 0.45). These two speakers realised aspirated Vf-transitions irrespective of the position of the Vf-target in phrase or of the vowel height involved.

Overall, this shows that the presence of aspiration in Vf-transitions is a matter of individual variation among people with the "same" phonological opposition, and not just a sex-related difference. We will also consider the nature of the interspeaker differences, but, briefly, it seems these two speakers have systematised their phonetic input in a way rather different to other speakers in this study. More speculatively, it might be possible to interpret such subtle interspeaker differences in production as being small potential steps along the road of systematic re-phonologisation, whether or not such non-normative variation ever achieves a sociolinguistic function or indeed results in diachronic change at a gross language level.

3.2 Vowel height

Across all speakers aspirated transitions of longer than 30ms occur more frequently in open vowels compared to more close vowels. Similar effects vowel height effects have been found in preaspirated voiceless stops in Sienese Italian (Hajek & Stevens, 2004).

One explanation to this pattern could be by accounting for the differences in intrinsic durations of different vowel heights as proposed by other authors (Hajek et al., 2004; Helgason, 2002) . We have indeed seen a significant correlation between vowel duration (including preaspiration) and vowel height in Section 2.4.2.2. However, the correlation was not very strong. Therefore, there is no one-to-one relationship reflecting preaspiration in this multidimensional phonetic space, which is clearly also shaped by aspiration-related parameters such as zero-crossing rate.

Note, however, that close vowels /i/ and /u/ in SSE show very large durational differences before /-voice/ and /+voice/ fricatives due to the Scottish Vowel Length Rule, while this is not the case for non-close vowel monophthongs (Aitken, 1981; Scobbie, Hewlett, & Turk, 1999). The fully phonated close vowel before /z/ might be twice as long as it is before /s/, and without observable preaspiration. The preaspiration after non-close vowels may, therefore, be an alternative way of supporting a broader (and more abstract) short/long durational difference across the vowel system.

Alternatively or additionally, the shorter (and less frequent) preaspiration before close vowels could be influenced by the presence of supraglottal constriction. Narrow supraglottal constriction for close vowels can cause a decrease in glottal friction (Stevens, 1998, p. 441 - 445): i.e. the constriction causes a reduction in the airflow compared to less constricted supraglottal configurations for more open vowels. The reduction in airflow subsequently causes a drop in transglottal pressure, and forces the vocal folds to abduct earlier. Despite shorter duration of preaspiration in close vowels, the supraglottal [ç]-like friction results in noise around F2 and F3 which is stronger in amplitude compared to that in more open vowels with less supraglottal constriction (Stevens, 1998), and might still be perceptually salient despite its brevity. Such patterns of vowel-dependent supraglottal friction have been found elsewhere, e.g. in preaspirated Faroese stops (Helgason, 2002), where such preaspiration is ‘normative’ (obligatory).

The interdependence of preaspiration and vowel height could also be explained by the forward masking effects known in human hearing. More salient (in terms of intrinsic spectral intensities) open vowels can produce greater masking of low intensity preaspiration which is otherwise similar to the preceding vowel in terms of resonances (Bladon, 1986). Therefore, longer preaspiration is needed to compensate for the greater masking effect, if we assume that preaspiration is intended by speakers to be perceptible.

3.3 Phrasal position

Phrase-final position of the Vf targets is found to condition the occurrence of preaspiration confirming our initial observations in Gordeeva & Scobbie (2004). This conditioning is significant in both subsets of data (male and female).

Previous research showed that preaspiration can be seen as a result of dissociation of glottal abduction and lingual gestures, with the substantially earlier onset of glottal opening relative to the oral stricture (Gobl & Ní Chasaide, 1988; Ní Chasaide & Gobl, 1993; Gobl et al., 1999). Preaspiration smears the temporal disjunction between the full vowel and the following consonant, dislocating its segmental homogeneity. In the light of the results here, it seems that the dissociation of laryngeal and supralaryngeal gestures appears to be the greatest at domain-final prosodic edges, i.e. pre-pausally.

The higher rate of appearance of transitional preaspiration at domain final prosodic edges is in line with previous findings that: (1) domain-final VC gestures are associated with increased articulatory strength, greater co-articulatory resistance, and longer and larger lingual articulator movements (Fougeron et al., 1997; Cho, 2001); (2) domain-final edges and accented targets have larger inter-gestural dissociation aimed to decrease segmental overlap (Edwards et al., 1991). Such segmentally-based domain-final strengthening carries a syntagmatic function in organizing speech flow into relevant prosodic domains and, thus, mediates speech communication. This supports the idea that prosodic constituency influences segmental processing in the mental representation (Keating et al., 2002).

The current study has its limitations in looking at a limited number of prosodic contexts, and only provides non-parametric rather than acoustic evidence of preaspiration discriminating between those different prosodic contexts. However, we can speculate that functionally, it could be that the speaker signals prosodic structure and phrasal accent via articulatory strengthening by dissociating the full vowel from the following voiceless fricative with a period of substantial aspiration, and the listener could make use of such an articulatory signature as a cue to a pre-pausal phrasal boundary, perhaps even interpreting it as a signal for turn-taking. Such a syntagmatic function of preaspiration remains to be investigated.

3.4 Enhancement of the fricative /voice/ contrast

An important finding in this study is that SSE fricative /voice/ contrast in phrase-final singleton targets is primarily encoded by Vf transitional mid- and high-frequency noise (measured as zero-crossing rate per sec, ZCR). ZCR, as specifically adapted in this study, reflects the presence/absence of aspiration noise in the transitions from vowels to /±voice/ fricatives. Additionally, the importance of ZCR as correlate of phonological /voice/ mirrors the importance of this parameter as an acoustic correlate of preaspiration shown in section 2.4.3. Therefore, we conclude that preaspiration serves to encode the SSE /voice/ contrast phrase-finally.

As aspiration-related transitional parameters have not been considered previously as a correlate of /voice/ in English fricatives, this characteristic may be variety-specific to SSE. It possibly plays a similar role in other preaspirating British English varieties like Middlesbrough English (Jones et al., 2003) or any other English variety with this characteristic. This implies that this optional characteristic is learnt, rather than automatic.

Although aspiration-related transitional parameters have not been addressed before as a correlate of fricative /voice/, it is possible that the extent of dissociation of laryngeal and oral stricture in previously studied English varieties (Non SSE British English: Haggard, 1978; Docherty, 1992; American English: Smith, 1997) was present but less substantial than in SSE, so was not noticed. The traditionally studied acoustic parameters of phrase-final fricative /voice/ in various English varieties, such as voicing offset, vowel and consonantal duration considered in previous research (Haggard, 1978; Docherty, 1992; Smith, 1997) also play a role in SSE , and contribute to the massive 92.2 % correct LDA classification of /voice/. However, the contribution of these parameters is less important than the high-frequency transitional aspiration reflected in ZCR.

This study, therefore, supports the abstractness and non-neutralizing nature of phonological /voice/ in English in general, such that it may be reflected in a number of acoustic correlates with no one-to-one mappings. In this multidimensional acoustic space of the various correlates can dynamically adjust and change in importance depending on phonetic structures and prosodic contexts involved.

Studies of fricative identification in fV sequences have shown that listener's attention can be shifted to formant transitions in contexts where static spectral fricative cues become insufficient. Wagner *et al* (2006) forwarded a hypothesis that "listener's attention to formant transitions for fricative identification is language-specific, and it is modulated by the presence of perceptually similar fricatives(e.g. /f θ/ in Spanish and English in the native phoneme inventory (Wagner, Ernestus, & Cutler, 2006, p.2268). Although the segmental context of preaspiration in Vf-transitions is prone to masking effects from the human auditory system which negatively influences its perception (Bladon, 1986), phonological systems that allow /VhC/ phonotactics seem to improve speaker's perception of such sequences (Mielke, 2003) in a language (and variety) specific way.

The extent and primacy of transitional aspiration-related correlates in phrase-final Vf-transitions in SSE is compatible with the above reasoning about the potential importance of such segment transitions (be it fV or Vf) and its learnability.

Native language learners' attention to Vf-transitions maybe mediated by the completeness of phonological neutralisation of /voice/ in word-final fricatives. Some British English varieties have partial phonetic devoicing of pre-pausal word-final /+voice/ obstruents, which is phonetically gradual without neutralizing the phonological contrast (Docherty, 1992). If there is pre-pausal devoicing without neutralisation, the /voice/ contrast must be being maintained somehow, by other phonetic correlates. Therefore, in a pre-pausal context, where important cues such as voicing offset and duration are demoted, transitional cues like preaspiration (very large timing dissociation between linguo-laryngeal gestures and/or very wide laryngeal abduction) may become promoted as more important in that specific context. In languages like Russian, Dutch or German where phonological neutralisation of /voice/ in phrase-final contexts has diachronically become complete, there is no (need for) such promotion of alternative transitional cues, and there are no reports of preaspiration for these languages. Similarly, in English varieties where final /+voice/ obstruents are not strongly devoiced phonetically, preaspiration for /-voice/ obstruents is probably less likely. Gestural dissociation as conditioned by prosodic context would, therefore, vary dialectally as a function of the phonological system, and the preservation of the /voice/ contrast in turn may be dependent on socially-mediated spread of the functional dissociation of laryngeal and supralaryngeal information.

This interpretation of the data is also compatible with the view that large preaspiration of voiceless fricatives is a learnt variety-specific characteristic in Scottish English and not universal (and pertaining cross-linguistically) as suggested in Gobl & Ní Chasaide (1999). Possibly, the occurrence of preaspiration may be mediated by the automatic ease to dissociate the laryngeal and oral stricture gestures before voiceless fricatives with the laryngeal abduction cross-linguistically preceding the lingual stricture (Gobl et al., 1999), since both preaspiration and oral frication requiring very large glottal opening (Löfqvist et al., 1987; Hoole, 1999). Although, surprisingly, there are no reports of preaspiration of voiceless fricatives of such large extent as here for SSE or Middlesbrough English (Jones et al., 2003) , it remains to be proven that more tightly-timed association of glottal abduction and oral constriction gestures in the Vf-transitions are possible in other languages or English varieties to provide unambiguous prove that extensive preaspiration reported in this study is an intentional target for Scottish English and, thus, is learnt.

3.5 Individual vs. variety-specific phonological systems

There seem to be different relationships among speakers in this paper in multidimensional phonetic correlates of fricative /voice/. While the SSE speakers seem to share similar phonologies in terms of maintaining /voice/, different speakers use different correlates of /voice/ available in their inventory. For example, the /voice/ contrast of speaker M5 is mainly cued by the aspiration-related parameters. While other speakers also employed voicing offset, and segmental duration on top of primary aspiration-related ZCR (see Figure 9), speaker M5 did not employ durational characteristics to cue /voice/. The relationship between the multidimensional phonetic characteristics in this speaker is shown in Figure 10.

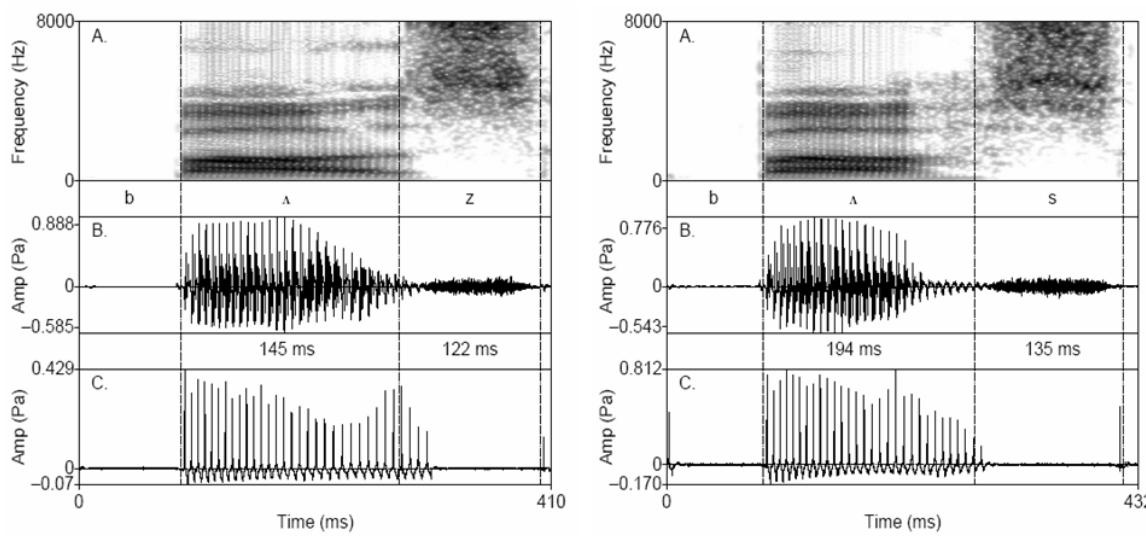


Figure 10. Example of /voice/ contrast for subject M5, producing “buzz” (left pane) and “bus” (right pane) in the same phrasal position and very similar speech rate. Pane A represents the spectrogram; Pane B represents acoustic wave; Pane C represents the time derivative of EGG lx waveform.

Figure 10 represents instances of “bus” and “buzz” produced by M5 in the same prosodic context and at quite similar speech rate. Unlike in other speakers, the duration of [ʌ] in “bus” is longer than in “buzz”. This relationship is mirrored in Speaker M5’s durational means in Figure 9. This individual relationship in vowel duration and ‘voicing effect’ clashes with the hypothesis of universal lengthening of vowels triggered by the following voiced oral obstruents (Chen, 1970). The consonantal duration differs by only by 13 ms, which is not found a perceptually relevant limens for fricative sounds (Jongman, 1989). While both /±voice/ fricatives are phonetically voiceless nearly throughout their durations, the voicing offset in time-derived lx domain of the EGG signal (Pane 3) differs only by 15 ms between the two instances in Figure 10. The only substantial difference in fricative /voice/ appears on Pane 1 in the sound spectrogram: i.e. the 64 ms long high frequency noise in the vowel-voiceless fricative transition on the right upper Pane, while the phonetic voicing persists, and there is no such abrupt interruption by noise in the "buzz" Vf-transition.

Figures 10 and 9 show that for a number of the acoustic parameters, the range of phonetic space used by the speakers as a whole for /+voice/ overlaps with the range used for /-voice/, so that a particular point on the range might encode either value of /+voice/, in a way which is speaker-dependent: hence the meaning of the feature /voice/ is in some regards arbitrary, gradient, and non-universal phonetically. Nevertheless, the speakers are remarkably consistent in the ways in which they offset their own individual acoustic properties of /±voice/, as shown the parallel nature of the lines in Figure 9. Therefore, the systematic relationship between /s/ and /z/ remains relatively constant, and non-arbitrary, even if in transcriptional terms we would have to represent it as ranging from [ʰs] vs. [s] to [s] vs. [z̥].

Such patterns resemble those shown for postaspiration in Shetlandic English (Scobbie, 2006), where the individual VOT patterns for /ptk/ were shown to function indexically ranging gradiently from short to long lag, and where there seemed to be a functional preservation of contrast between /p/ and /b/, such that speakers with increasingly shorter-lag /p/ tended to have more prevoiced /b/.

4. Conclusions

This study examined the phonetic conditioning and function of aspiration as an acoustic event during vowel/voiceless fricative transitions in Scottish Standard English. Though non-obligatory and gradient, the extent of this characteristic is sufficiently large to merit the use of the phonetic label 'preaspiration' with respect to fricatives as previously conceived for other obstruents (Laver, 1994; Ladefoged et al., 1996; Helgason, 2002).

General lack of crosslinguistic reports suggests that preaspiration it is not just an automatic aerodynamic consequence of vowel-voiceless fricative production as previously thought (Gobl et al., 1999), but can also be a *variety-specific* optional characteristic resulting from a *learnt* dissociation of linguo-laryngeal stricture gestures in anticipation of voiceless fricatives, in a way similar to that observed in stops (Gobl et al., 1988; Ní Chasaide et al., 1993; Gobl et al., 1999).

There is structural phonetic variation in the distribution of preaspirated fricatives in SSE: its extent varies depending on vowel height and phrasal position. With reference to the latter factor, preaspiration may accord with the 'contrast maximisation principle' (Cho, 2001) by enhancing syntagmatic phonological contrasts in targets under sentence accent at domain-final edges assisting prosodic structure and speech communication; as well as by enhancing paradigmatic fricative /±voice/ contrast in phrasal positions, especially pre-pausally, where 'normal' phonation and voicing are less reliable cues. Therefore, SSE preaspiration may serve to simultaneously enhance cues to prosodic constituency and phonemic contrasts in /voice/ and stop-fricative manner.

Both male and female speakers equally produce preaspirated transitions in terms of frequency of occurrence, although female speakers produce preaspiration longer in duration, reflecting physiological and/or sociolinguistic differences. In multidimensional acoustic terms, individual differences are found to be more important than sex-related differences. From one point of view, individual speakers span a wide range of productions that argue for an abstract and partially arbitrary interpretation of abstract phonological labels like /voice/, because intermediate tokens can only be assigned to a particular phoneme if the speaker's identity (and the prosodic context) are known. On the other hand, each speaker's opposition between /-voice/ and /+voice/ seems remarkably consistent in maintaining the contrast, as shown by the comparable jumps made by different speakers for final fricative /±voice/ conditions along a number of acoustic parameters related to voicing offset, segmental duration and voice quality (aspiration). One of these voice quality parameters: i.e. the timing differences in the relative onset of a major increase in zero-crossing rate in Vf-transitions (reflecting aspiration noise present in mid/high spectral frequencies) – is found to be more consistent in this study than the more traditional phonetic correlates of phrase-final fricative /voice/.

Acknowledgements

We would like to thank the UK Economics and Social Research Council (ESRC grant PTA-026-27-0368 to Gordeeva) for funding this research. Many thanks to Tine Mooshamer for her invaluable advice about the use of acoustic measures in this paper, and to Steve Cowen for technical assistance. Thanks also to our subjects, and to delegates at BAAP 2004.

References

Agutter, A. (1988). The non-so-Scottish Vowel Length Rule. In J.M.Anderson & N. MacLeod (Eds.), *Edinburgh Studies in the English Language* (pp. 120-132). Edinburgh: John Donald Publishers.

Aitken, A. J. (1981). The Scottish Vowel Length Rule. In M.L.Benskin & M. Samuels (Eds.), *So Many People, Longages, and Tongues, Edinburgh: Middle English Dialect Project* (pp. 131-157).

Balise, R. R. & Diehl, R. L. (1994). Some distributional facts about fricatives and a perceptual explanation. *Phonetica, 51,* 99-110.

Bladon, A. (1986). Phonetics for hearers. In G.McGregor (Ed.), *Language for Hearers* (pp. 1-24). Oxford, U.K.: Pergamon.

Boersma, P. & Weenink, D. (2006). PRAAT, a system for doing phonetics by computer. www.praat.org, *version 4.3.04*.

Browman, C. P. & Goldstein, L. (1992). Articulatory Phonology: An overview. *Phonetica, 49,* 155-180.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22,* 129-159.

Chirrey, D. (1999). Edinburgh: descriptive material. In P.Foulkes & G.Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp. 223-229). London, UK: Arnold.

Cho, T. (2001). Effects of prosody on articulation in English. *UCLA Dissertation in Linguistics,* 1-202.

Docherty, G. J. (1992). *The timing of voicing in British English Obstruents*. Berlin: Netherlands Phonetics Archives, 9, Foris.

Docherty, G. J. & Foulkes, P. (1999). Derby and Newcastle: instrumental phonetics and variationist studies. In P.Foulkes & G.Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp. 47-71). London, UK: Arnold.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics lengthening. *Journal of Acoustical Society of America, 89,* 369-382.

Fant, G., Kruckenberg, A., & Nord, L. (1991). Prosodic and segmental speaker variations. *Speech Communication, 10,* 521-531.

Fougeron, C. & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America, 101,* 3728-3740.

Gobl, C. & Ní Chasaide, A. (1988). The effects of adjacent voiced/voiceless consonants on the vowel voice source: a cross-language study. *STL-QPSR, 2-3.*

Gobl, C. & Ní Chasaide, A. (1999). Voice source variation in the vowel as a function of consonantal context. In W.J.Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data, Techniques* (pp. 122-143). Cambridge University Press.

Gordeeva, O. B. (2005). *Language interaction in bilingual acquisition of sound structure: a longitudinal study of vowel quality, duration and vocal effort in pre-school children speaking Scottish English and Russian*. Edinburgh: Unpublished Ph.D. thesis. Queen Margaret University College.

Gordeeva, O. B. & Scobbie, J. M. (2006). *Phonetic conditioning of word-final ejective stops in the speech of Scottish English pre-school children*. Edinburgh: Colloquium of the British Association of Academic Phoneticians. Queen Margaret University College.

Gordeeva, O. B. & Scobbie, J. M. (2004). *Non-normative preaspiration of voiceless fricatives in Scottish English*. Cambridge: Colloquium of the British Association of Academic Phoneticians. University of Cambridge.

Haggard, M. (1978). The devoicing of voiced fricatives. *Journal of Phonetics, 6,* 95-102.

Hajek, J. & Stevens, M. (2004). Mechanisms of sound change in Romance: From gemination to degemination in Italy. *Paper presented at LabPhon IX, Univ.of Illinois atUrbana-Champaign, July 7-9..*

Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Journal of Acoustical Society of America, 101,* 466-481.

Helgason, P. (2002). *Preaspiration in the Nordic Languages: synchronic and diachronic aspects*. Unpublished Ph.D. Thesis. Department of Linguistics, Stockholm University.

Henton, C. G. & Bladon, R. A. W. (1985). Breathiness in normal female speech: inefficiency versus desirability. *Language and Speech, 5,* 221-227.

Hillenbrand, J., Cleveland Ronald A, & Ericson Robert L (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research, 37,* 769-778.

Holmberg, E. B., Hillman Robert E, Perkell Joseph S, Guiod Peter C, & Goldman Susan L (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research, 38,* 1212-1223.

Hoole, P. (1999). Techniques for investigating laryngeal articulation. Section A: Investigation of the devoicing gesture. In W.J.Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data, Techniques* (pp. 294-300). Cambridge University Press.

Jones, M. J. & Llamas, C. (2003). Fricated pre-aspirated /t/ in Middlesbrough English: an acoustic study. In *Proceedings of the 15th International Congress of Phonetic Sciences* (4 ed., pp. 655-658). Barcelona.

Jongman, A. (1989). Duration of frication noise required for identification of English fricatives. *Journal of Acoustical Society of America, 85,* 1718-1725.

Keating, P. & Shattuck-Hufnagel, S. (2002). A Prosodic View of Word Form Encoding for Speech Production. *UCLA Working Papers in Phonetics,* 112-157.

Kim, C.-W. (1970). A theory of aspiration. *Phonetica, 21,* 107-116.

Klatt, D. H. & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of Acoustical Society of America, 87,* 820-857.

Koenig, L. L., Mencl, W. E., & Lucero, J. C. (2005). Multidimensional analysis of voicing offsets and onsets in female subjects. *Journal of Acoustical Society of America, 118,* 2535-2550.

Ladefoged, P., Ladefoged, J., Turk, A., Hind, K., & Skilton St.John (1999). Phonetic structures of Scottish Gaelic. *Journal of the International Phonetic Association, 28,* 1-41.

Ladefoged, P. & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford, UK: Blackwell Publishers.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*.

Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word, 20,* 384-422.

Löfqvist, A. (1995). Laryngeal mechanisms and interarticulator timing in voiceless consonant production. In F.Bell-Berti & L. J. Raphael (Eds.), *Producing Speech: Contemporary Issues. For Katherine Safford Harris* (pp. 99-116). Woodbury, NY: AIP Press.

Löfqvist, A., Koenig, L. L., & McGowan, R. S. (1995). Vocal tract aerodynamics in /aCa/ utterances: Measurements. *Speech Communication, 16,* 49-66.

Löfqvist, A. & McGarr, N. S. (1987). Laryngeal dynamics in voiceless consonant production. In T.Baer, C. Sasaki, & K. S. Harris (Eds.), *Laryngeal function in phonation and respiration* (pp. 391-402). Boston: College Hill.

Löfqvist, A. & Yoshioka, H. (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of Acoustical Society of America, 68,* 792-801.

Marasek, K. (1997). *Electroglottographic description of voice quality*. Stuttgart: Habilitationschift Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

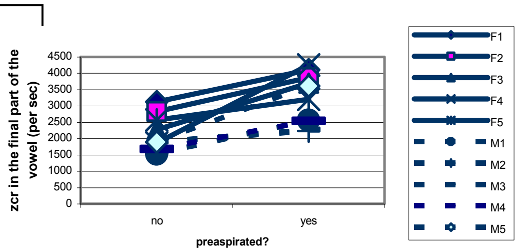
McKenna, G. (1988). *Vowel Duration in the Standard English of Scotland*. Edinburgh: unpublished MSc thesis, University of Edinburgh.

Mielke, J. (2003). The Interplay of Speech Perception and Phonology: Experimental Evidence from Turkish. *Phonetica, 60,* 208-229.

Murphy, P. J. (1999). Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis. *Journal of Acoustical Society of America, 105,* 2866-2881.

Ní Chasaide, A. (1987). Glottal control of aspiration and of voicelessness. In *Proceedings of the 11th International Conference of Phonetic Science* (pp. 28-31). Tallinn, Estonia.

Ní Chasaide, A. & Gobl, C. (1993). Contextual variation of the vowel voice source as a function of adjacent consonants. *Language and Speech, 36,* 303-330.



of sound change. In C.Jones (Ed.), *Historical*
d Perspectives (pp. 237-278). London: Longman.
C.Rowden (Ed.), *Speech Processing* (pp. 35-73).
ook Company.

Schaeffler, F. (2005). *Phonological Quantity in Swedish Dialects: Typological aspects, phonetic variation and diachronic change*. (Phonum ed.) (vols. 10) Umeå University, Sweden.

Scobbie, J. M. (2006). Flexibility in the face of incompatible English VOT systems. In L.M.Goldstein, C. Best, & D. Whalen (Eds.), *Papers in Laboratory Phonology 8: Varieties of Phonological Competence* (pp. 367-392). Berlin: Mouton de Gruyter.

Scobbie, J. M., Hewlett, N., & Turk, A. (1999). Standard English in Edinburgh and Glasgow: the Scottish Vowel Length Rule revealed. In P.Foulkes & G.Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp. 230-245). London: Arnold.

Silverman, D. (2003). On the rarity of pre-aspirated stops. *Journal of Linguistics, 39,* 575-598.

Smith, C. L. (1997). The devoicing of /z/ in American English: effects of local and prosodic context. *Journal of Phonetics, 25,* 471-500.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, Massachusetts: The MIT Press.

Stuart-Smith, J. (2004). The phonology of Scottish English. In Kortman B & Schneider E.W (Eds.), *A Handbook of Varieties of English* (pp. 47-67). Berlin: Mouton de Gruyter.

Stuart-Smith, J. (1999). Voice Quality in Glaswegian. In J.J.Ohala, Y. Hasegawa, Ohala M, D. Granville, & A. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 2553-2556). San Francisco: University of California.

Thráinsson, H. (1979). *On Complementation in Icelandic*. New York: Garland Publishing Inc.

van Dommelen, W. A. (1998). Production and perception of preaspiration in Norwegian. In P.Branderud & H. Traunmüller (Eds.), *Proceedings of FONETIK 98* (pp. 20-23). Stockholm: Stockholm University.

Vaux, B. (1998). The laryngeal specification of fricatives. *Linguistic Inquiry, 29,* 497-511.

Wagner, A., Ernestus, M., & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *Journal of Acoustical Society of America, 120,* 2267-2277.

Watt, D. & Allen, W. (2003). Tyneside English. *Journal of the International Phonetic Association, 33,* 267-271.

Wells, J. (1982). *Accents of English*. (vols. 1-3) Cambridge: Cambridge University Press.

Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *Journal of Acoustical Society of America, 71,* 1544-1550.

.